NeurIPS Workshop on What Can('t) Transformers Do?

Transformers in the Dark: Navigating Unknown Search Spaces via Noisy Feedback

Jungtaek Kim[1] Ziqian Lin[1] Thomas Zeng[1] Minjae Lee[3] Chungpa Lee[4] Jy-yong Sohn[4] Hyung Il Koo[3] Kangwook Lee[1,2]

[1]University of Wisconsin–Madison [2]KRAFTON AI [3]FuriosaAI [4]Yonsei University

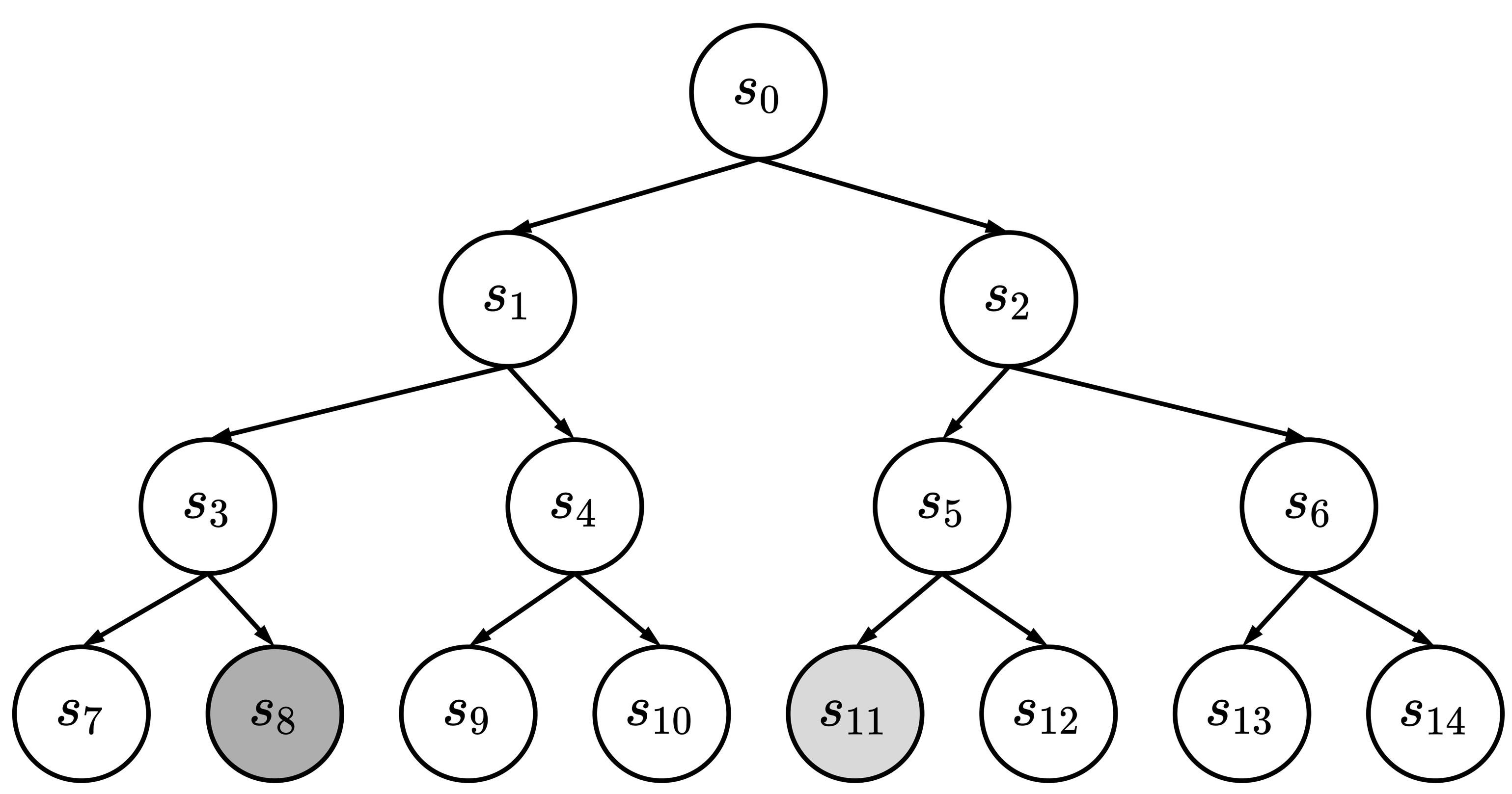
Motivation

- Effective problem solving often follows an iterative process of
  - generating diverse idea candidates,
  - selecting the most promising one to pursue,
  - evaluating its potential.
- Large Language Models (LLMs) suggest that these systems may implicitly implement such a process.
- This problem-solving process can be enhanced when LLMs are paired with external search algorithms.

Can LLMs or their underlying Transformer architectures fully internalize the search algorithm?

Problem Setup

- Each problem instance consists of a finite rooted search tree $\mathrm{T} = (S, N)$ with maximum depth D, where S is a state set and $N: S \to 2^S$ is a successor function.
- Also, $r: S \to [0, 1]$ is a bounded reward function.
- The goal states $S_{\mathrm{goal}} = s \in S : r(s) > 0$ are leaves.
- The tree T is hidden from the agent.
- We define the value $V(s)$ as a distribution that provides noisy estimates of the expected reward from state s.
- Under this setting, the agent interacts with the environment for $T \ll |S|$ steps.
- Starting from the root s_0, at each step t the environment reveals the children $N(s_t)$ and a value sample $s_t \sim V(s_t)$.
- The agent then selects the next state s_{t+1} from the frontier of unvisited children, according to its policy.



Reference Search Algorithms

- Uniform Leaf Sampling
- Greedy Leaf Sampling
- Uniform Path Sampling
- Policy-Guided Path Sampling (Pure Exploration, Greedy, and UCT)

Theoretical Analysis

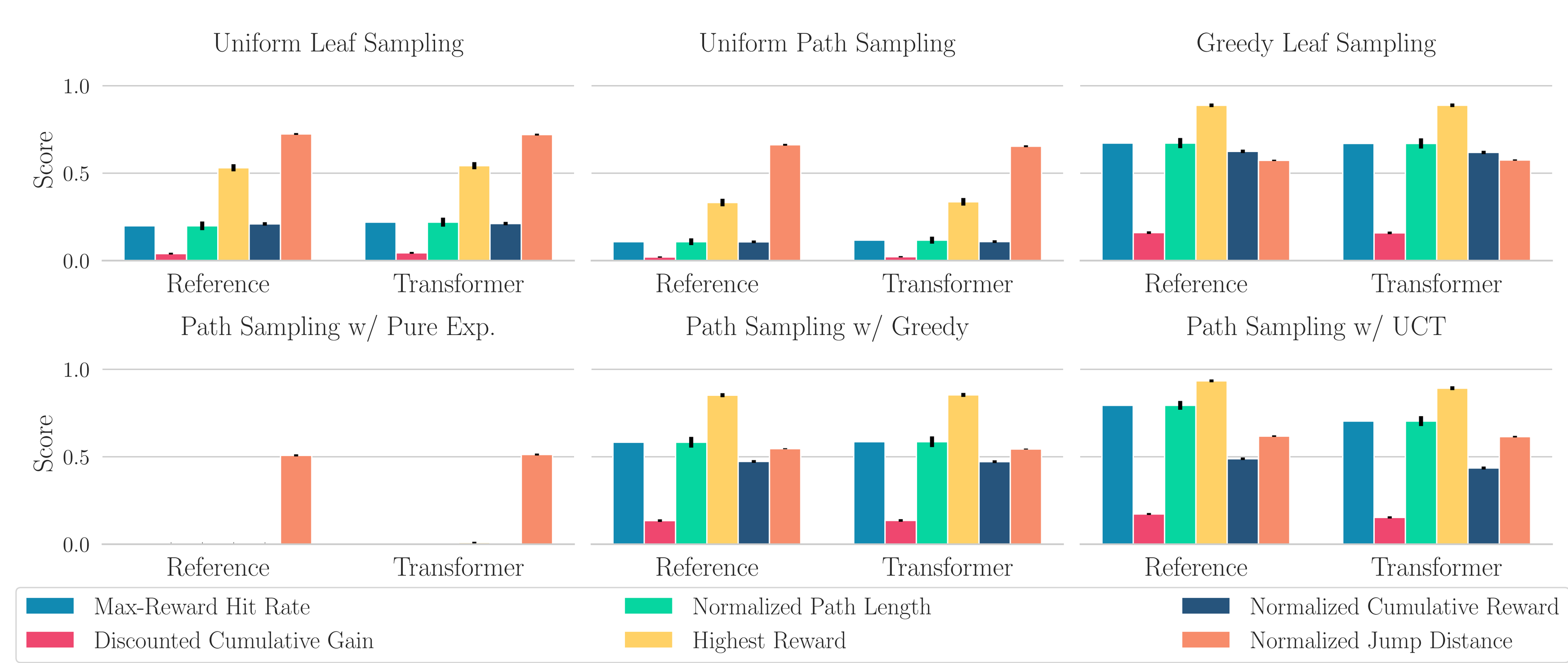
- We provide explicit weight constructions for Transformers with constant depth and embedding dimension linear in the search budget T and branching factor B.

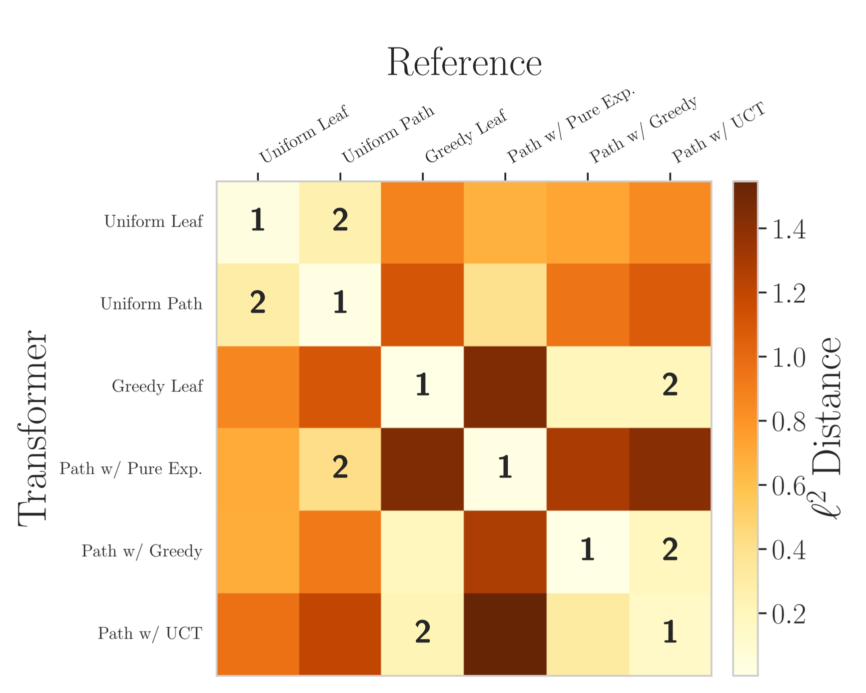
Theorem 1 *There exist 3-layer Transformers with embedding dimension $d = \Theta(BT)$ that exactly implements the uniform and greedy leaf sampling policies when using a sequential encoding of the trajectory using tokens linear in the number of search iterations.*

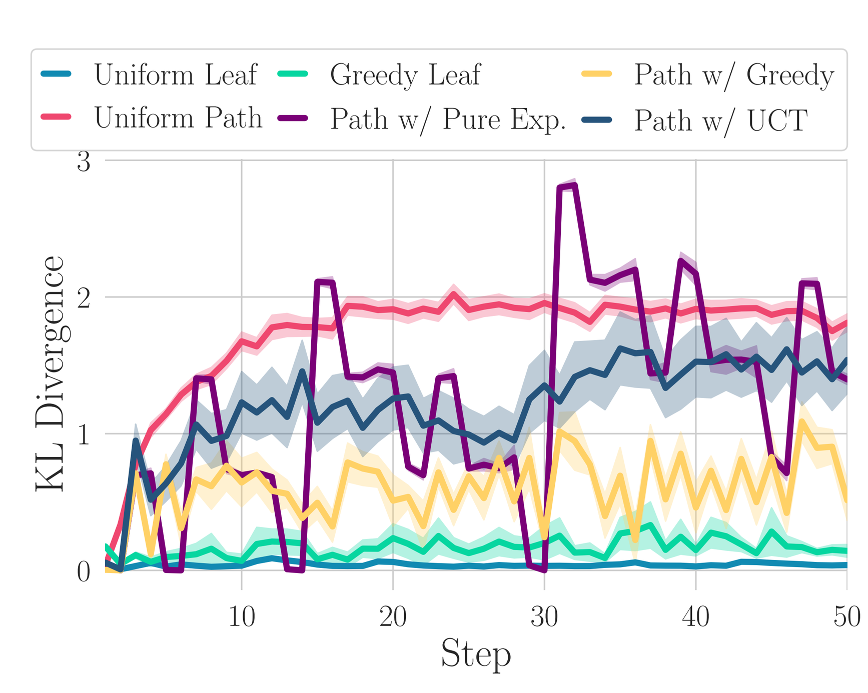
Theorem 2 *There exist 12-layer Transformers with embedding dimension $d = \Theta(BT)$ that exactly implements random, greedy and UCT based path-sampling policies when using a sequential encoding of the trajectory using tokens quadratic in the number of search iterations.*

Empirical Analysis

- We provide explicit weight constructions for Transformers with constant depth and embedding dimension linear in the search budget T and branching factor B.







Conclusion

- In this paper, we have introduced synthetic and fully controllable search benchmarks.
- Our theoretical and empirical analyses demonstrate that Transformers can effectively implement classical search strategies within these environments.

PDF

My Profile