Reading Between the Lines: Teaching Open-Source Embodied Agents to Resolve Vague Human Instructions

Souradeep Mukhopadhyay[1] Vedang Vasant Avaghade[1] Ali Payani[2]
Gaowen Liu[2] Jayanth Srinivasa[2] Chitta Baral[1]
[1]Arizona State University [2]Cisco Research
{smukho11, chitta}@asu.edu

Abstract

Open-source large language models (LLMs) suffer from catastrophic performance collapse (19–58% success rate) when faced with ambiguous, human-like instructions characterized by implicit pronouns and noisy dialogue. To bridge this gap, we introduce FAST-PLAN, a scalable, annotation-free framework that synthesizes training data by identifying and refining a model's own planning failures. By fine-tuning open source backbones (Llama-3.2-Inst., Qwen-3-4B-Inst.) on these 1.8K failure-driven trajectories, we achieve a +50 to +80 percentage points success rate increase on REI-Bench and demonstrate zero-shot generalization to the CLARA (+6.14 F1) and AmbiK, two out-of-distribution embodied ambiguity benchmarks.

1. Introduction

Embodied agents must navigate the "semantic gap" between vague human commands (e.g., "heat it and put it there") and deterministic robotic actions. While frontier models (GPT-4o) possess innate reasoning to resolve such ambiguity, they are impractical for on-device robotics due to API latency, cost, and privacy concerns. Conversely, open-source models lack the targeted training signals required for systematic disambiguation. We propose a paradigm shift from inference-time prompting to deep, failure-driven training supervision grounded in the AI2-THOR environment.

2. Methodology: The FAST-PLAN Pipeline

FAST PLAN generates training data through an automated failure driven pipeline that turns ambiguous household instructions into grounded planning supervision. First, the system samples valid task settings in AI2 THOR [3] and constructs each scenario using deterministic scene information, including the target object, destination, distractor objects, and expert action plan. An augmentor then rewrites each task as a natural human instruction with varying ambiguity, such as explicit object names, attribute based references, pronouns, short context, or noisy dialogue. A resolver predicts the intended object and destination from the dialogue, and a plan writer produces an executable robot plan from the resolved entities and scene description. The generated plan is passed through a grounding checker that removes physically invalid actions, nonexistent objects, and incomplete plans. Finally, a critic model audits the interaction, assigns fine grained failure labels, and provides corrective feedback, so only grounded successful plans and useful failure corrected trajectories are retained as supervised fine tuning data for robust ambiguity aware embodied planning. The overall 6 stage pipeline is shown in Fig. 1.

3. Experimental Results

Experimental results across three model scales demonstrate that FAST-PLAN effectively bridges the "semantic gap" where inference-time interventions falter. As shown in Fig. 2, open-source backbones collapse on REI-Bench [2] using standard prompting—peaking at 58.7% (Qwen-4B via ICL)—whereas our failure-driven SFT catapults success rates to 93.8–96.6%. This leap confirms that referential ambiguity is a training distribution deficit rather than a prompt-engineering bottleneck. Furthermore, Fig. 3 illustrates zero-shot generalization to OOD scenarios, yielding a +5.8% SR gain on AmbiK [1] safety tasks and a +6.14 macro-F1 increase on CLARA [4]. These results indicate that failure-aware synthetic trajectories enable compact models to match proprietary system performance for on-device deployment.

4. Conclusion

We presented FAST-PLAN, a scalable, annotation-free, failure-driven framework that transforms open-source LLM planning mistakes into high-quality supervised fine-tuning

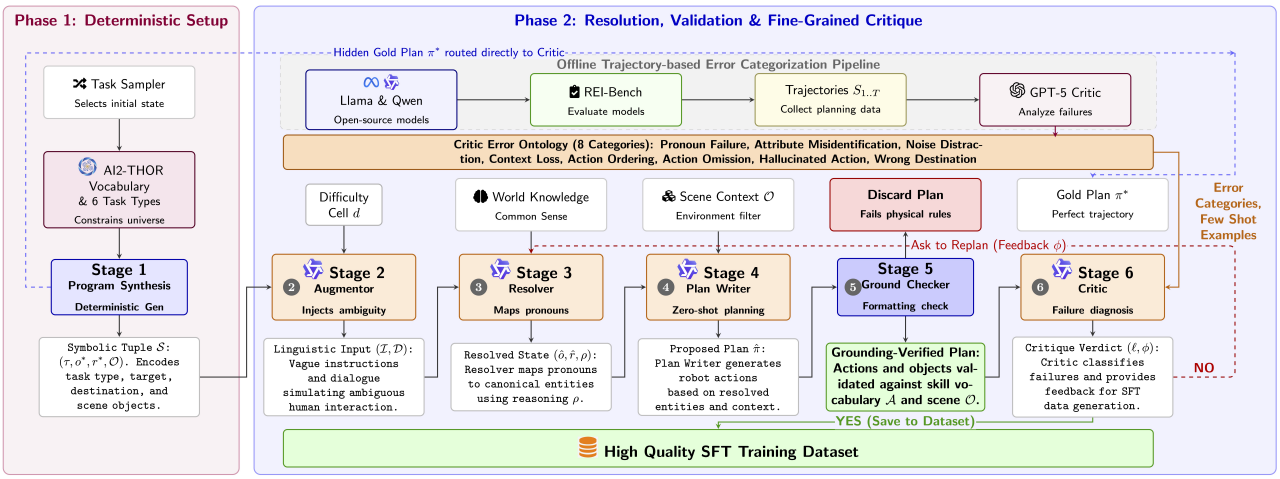

Figure 1. Overview of the FAST-PLAN data generation and validation pipeline. Phase 1 deterministically samples AI2-THOR task configurations and synthesizes symbolic task tuples, defining the target object, destination, scene context, and gold plan. Phase 2 converts these tuples into ambiguous linguistic instructions, resolves entities, generates candidate robot plans, and validates them through grounding checks and critic-based failure diagnosis. Invalid plans are rejected or sent for replanning with feedback, while grounding-verified plans are retained as high-quality SFT training data.

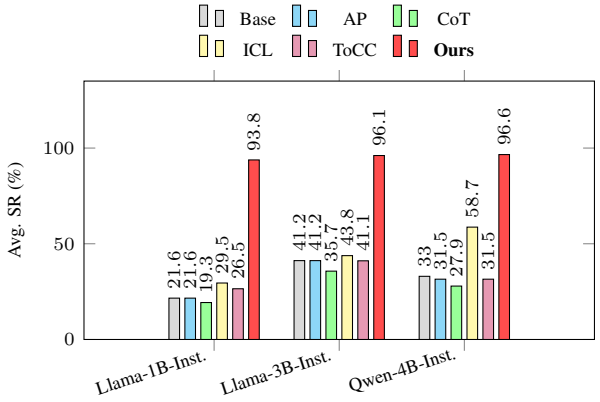

Figure 2. REI-Bench Success Rates. Our approach significantly outperforms all prompting baselines across models.

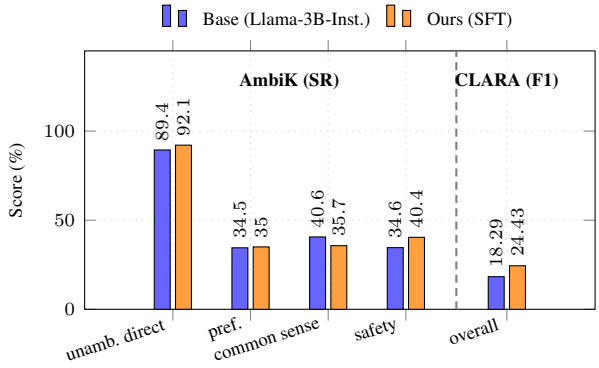

Figure 3. Out-of-domain results comparing base and SFT models. AmbiK denotes success rate (%); CLARA denotes macro-F1 (%).

signals Fine-tuning Llama-3.2-1B-Inst., Llama-3.2-3B-Inst., and Qwen3-4B-Inst. on synthesized trajectories yields 94 to 96% success rates on REI-Bench, achieving a 50 percentage point improvement over advanced prompting baselines Crucially, these models exhibit strong zero-shot transfer to CLARA (+6.14 macro-F1) and AmbiK (+5.8 on safety-critical categories), confirming that the improvements reflect genuine disambiguation competence. By eliminating dependence on frontier API-based models, FAST-PLAN enables privacy-preserving, on-device deployment without recurring API inference cost, lowering barriers to accessible robotic assistance These results establish that LLM fragility under human-like ambiguity is fundamentally a training data problem, and FAST-PLAN offers a practical solution

References

[1] Anastasiia Ivanova, Eva Bakaeva, Zoya Volovikova, Alexey K. Kovalev, and Aleksandr I. Panov. Ambik: Dataset of ambiguous tasks in kitchen environment. *ArXiv*, abs/2506.04089, 2025. 1

[2] Chenxi Jiang, Chuhao Zhou, and Jianfei Yang. Rei-bench: Can embodied agents understand vague human instructions in task planning? *ArXiv*, abs/2505.10872, 2025. 1

[3] Eric Kolve, Roozbeh Mottaghi, Winson Han, Eli VanderBilt, Luca Weihs, Alvaro Herrasti, Matt Deitke, Kiana Ehsani, Daniel Gordon, Yuke Zhu, Aniruddha Kembhavi, Abhinav Kumar Gupta, and Ali Farhadi. Ai2-thor: An interactive 3d environment for visual ai. *ArXiv*, abs/1712.05474, 2017. 1

[4] Jeongeun Park, Seungwon Lim, Joonhyung Lee, Sangbeom Park, Minsuk Chang, Youngjae Yu, and Sungjoon Choi. Clara: Classifying and disambiguating user commands for reliable interactive robotic agents. *IEEE Robotics and Automation Letters*, 9:1059–1066, 2023. 1