

CS839 Foundation Models: Presentation Proposal

Sujan Reddy Ale, Gaurav Batra, Arshad Aslam Kazi, Shivam Mittal, Grant Waldow

October 23, 2025

Presentation Paper

We propose to present: ”**RLVR-World: Training World Models with Reinforcement Learning**” by Jialong Wu, et al., to appear in NeurIPS 2025 [Wu et al., 2025].

Reason for Selection & Core Idea

We chose this paper because it tackles a fundamental limitation in current world model training: the mismatch between standard training objectives and task-specific evaluation goals. Traditional objectives like Maximum Likelihood Estimation (MLE) optimize for likelihood rather than for metrics that truly reflect model quality, such as prediction accuracy or perceptual realism.

This paper proposes **RLVR-World**, a post-training framework that applies **Reinforcement Learning with Verifiable Rewards (RLVR)** to directly optimize world models for these downstream metrics. Instead of relying on human feedback or proxy losses, RLVR uses automatically computable, deterministic rewards—such as correctness or perceptual quality—evaluated on decoded outputs. This allows models to improve their practical performance across diverse domains, including text games, web navigation, and robot manipulation.

Fit for Course

This paper is an ideal case study for this course, as it demonstrates:

- **Post-Training Alignment:** A powerful method to fix the limitations of standard pre-training objectives.
- **RLVR as an RLHF Alternative:** A key alignment technique for objective, verifiable tasks (e.g., robotics, web agents) where ”human preference” isn’t the main goal.
- **Models as Simulators:** The advanced use of foundation models as dynamic ”world models” for autonomous agents.
- **The Unified Transformer:** Applying a single post-training paradigm to radically different modalities (text and video).

References

Jialong Wu, Shaofeng Yin, Ningya Feng, and Mingsheng Long. RLVR-World: Training world models with reinforcement learning, 2025. URL <https://arxiv.org/abs/2505.13934>.