Constrained Multi-Objective Post-training for Vision-Language-Action Models

CODY HILL • 7/29/26

DR. YUE WANG & ZAIN FARHAT



National Science Foundation

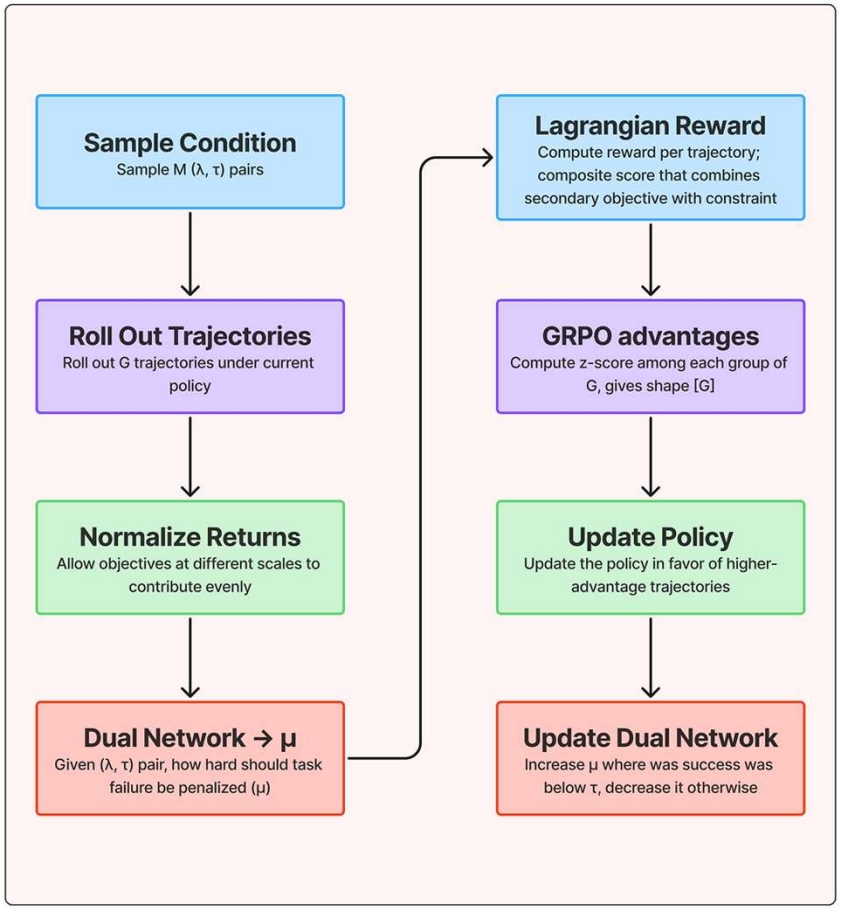


Center for Research in Computer Vision



MISSISSIPPI STATE UNIVERSITY

Problem Overview



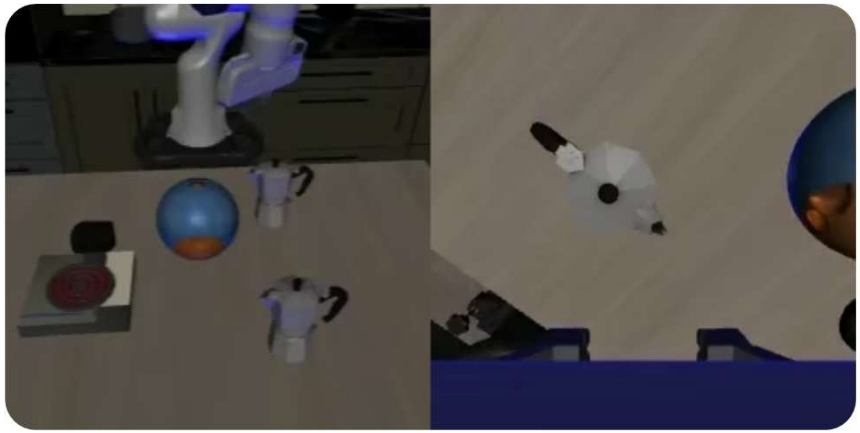
Task

- VLAs → Success only
- GRPO with task-success floor
- Optimized secondary objectives
- Policy π with maximum returns

Vision-Language-Action

- Qwen 0.5B
- Wrist image + 3rd person
- Predefined text inputs
- Output action distribution

"put both moka pots on the stove"



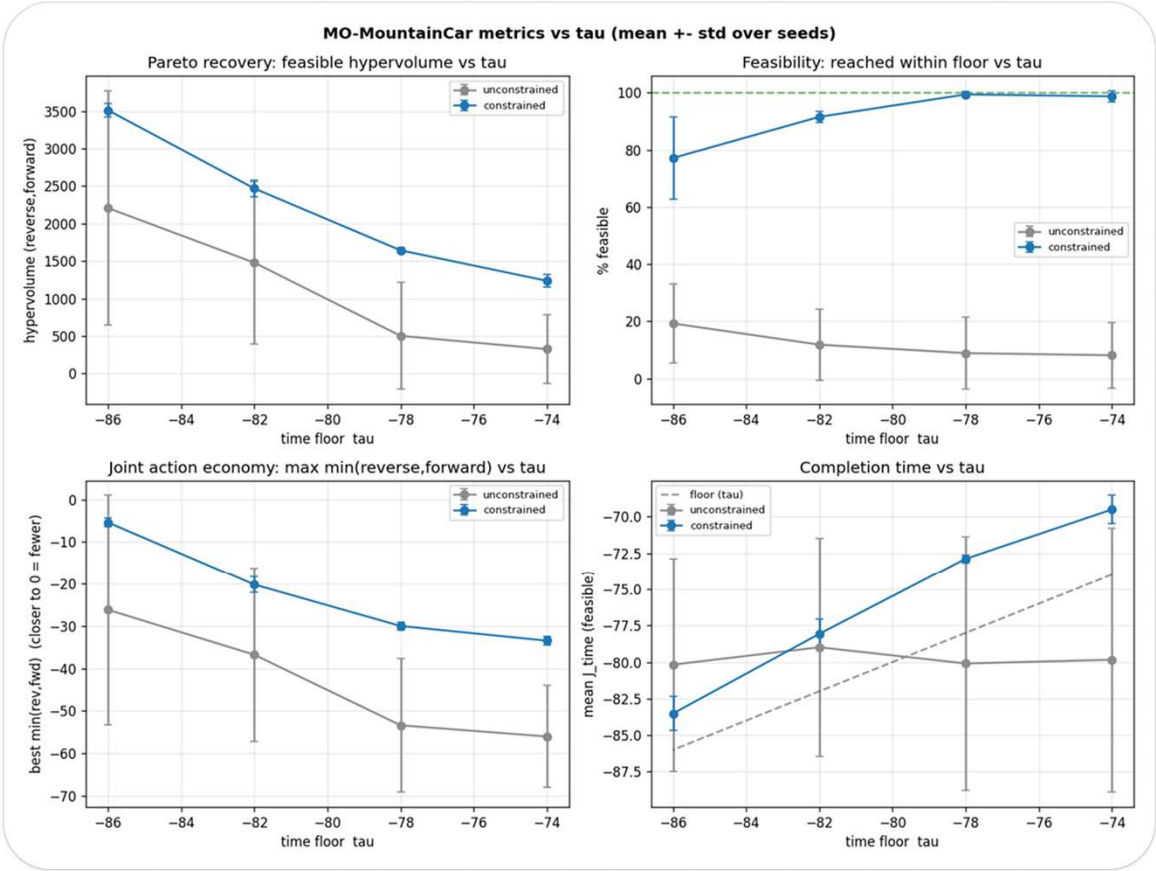
Constraint

	Bad	Feasible	Optimal
Success	0.64	0.85	0.85
Efficiency	0.84	0.41	0.77
Safety	0.73	0.51	0.68
Smoothness	0.75	0.43	0.75

$\tau = 0.8$

Progress: Toy Validation

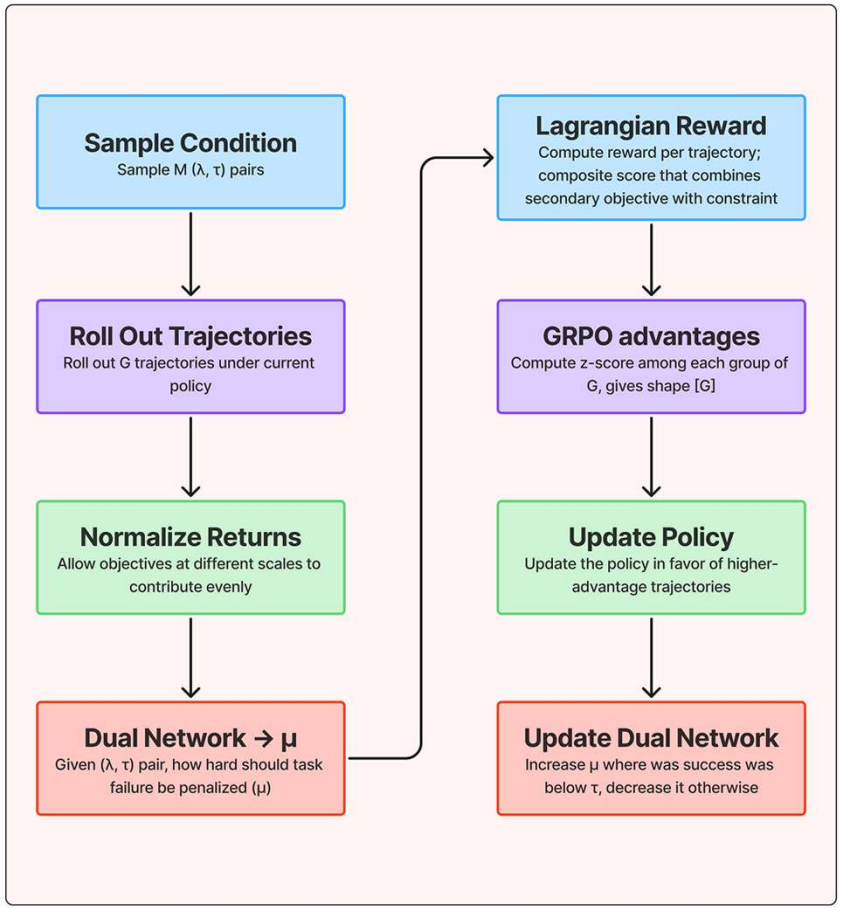
- Constraint response
- Preference response



Next Steps

Next Steps

- Measure tradeoffs in LIBERO
- Observe in the real world
- Publish



Constrained Multi-Objective Post-training for Vision-Language-Action Models

CODY HILL • 7/29/26

DR. YUE WANG & ZAIN FARHAT



National Science Foundation



Center for Research in Computer Vision



MISSISSIPPI STATE UNIVERSITY