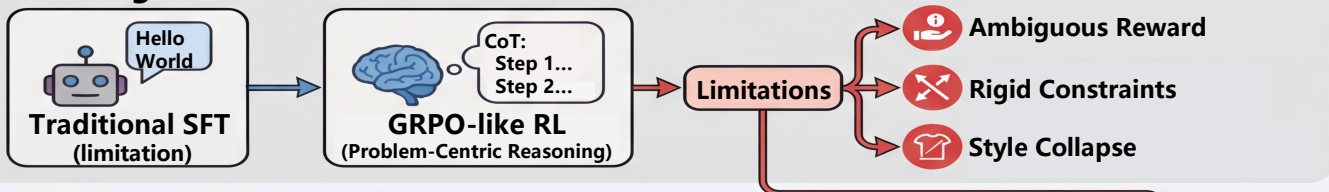


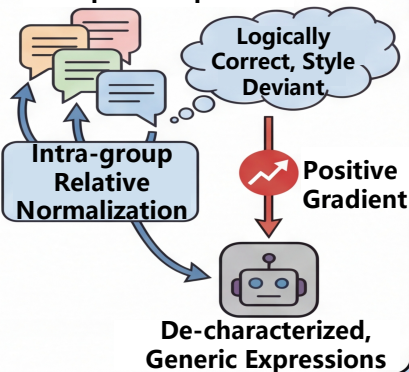
Paradigm Shift & GRPO Limitations



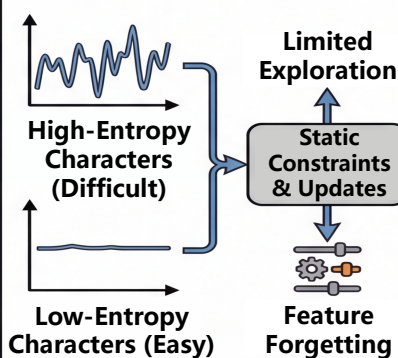
GRPO-like Method Limitations: Character-Centric Challenges

1. Ambiguous Reward Signal

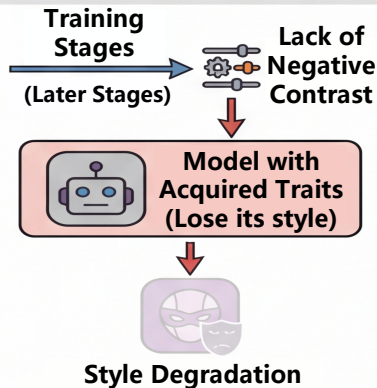
Sampled Responses



2. Rigid Optimization Constraints



3. Risk of Style Collapse



Outcome & Final Model



CRPO
Optimization Objective



Deep CoT with Character Fidelity & Consistency
Character-Centric RPA