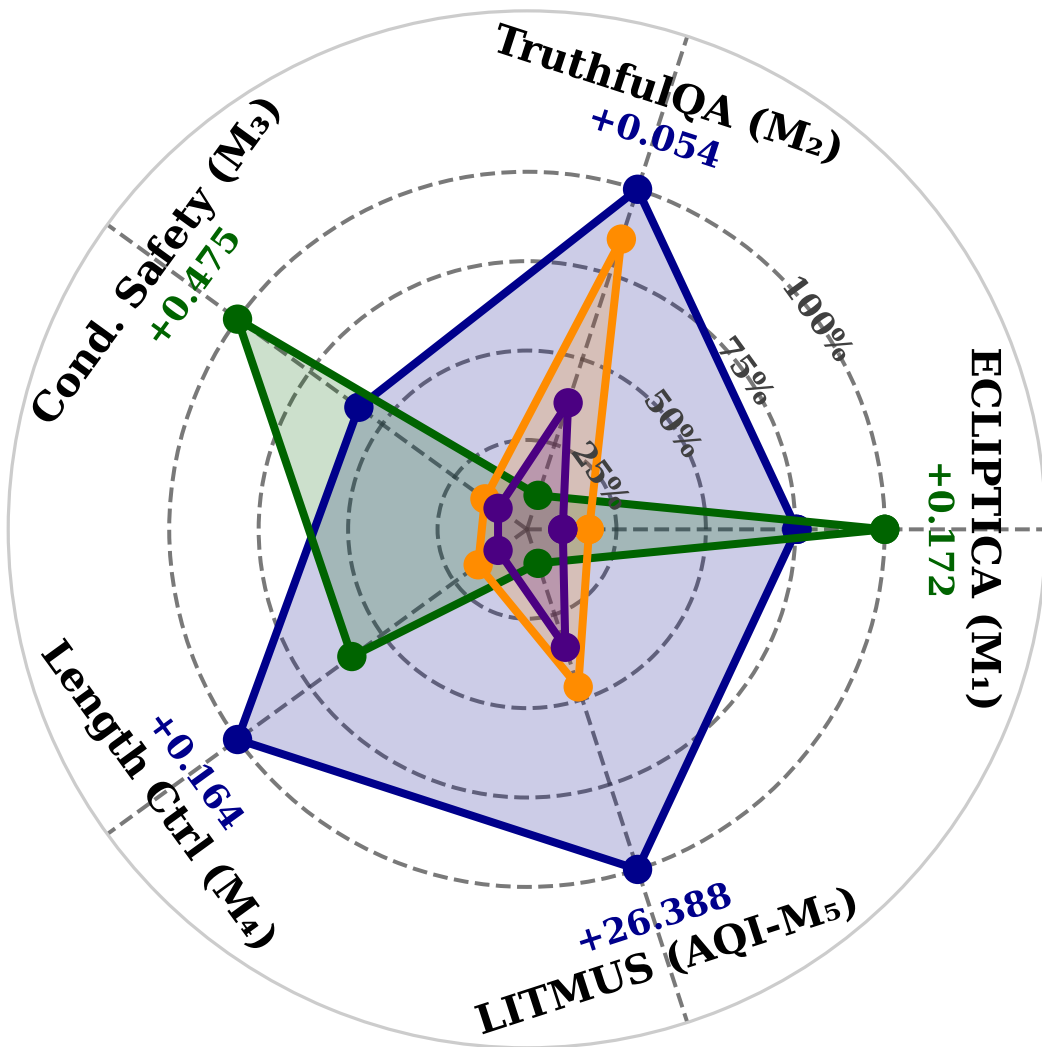


Instruction Alignment Efficiency (Llama-3.1-8B)

(Average Radius = Overall Performance)



Method (Avg %)

● CITA (86.7%)

● GRPO (36.1%)

● DPO (56.1%)

● PPO (20.4%)