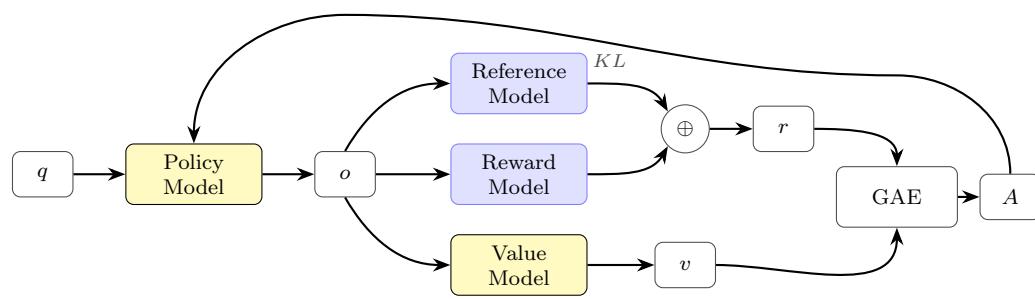


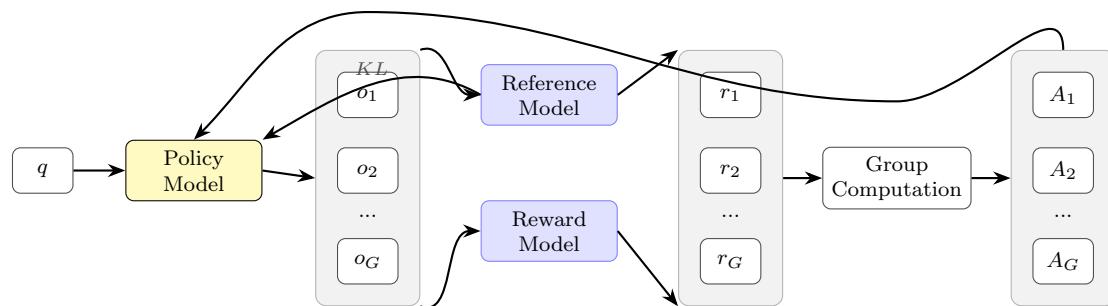
PPO



Trained Models

Frozen Models

GRPO



RLOO

