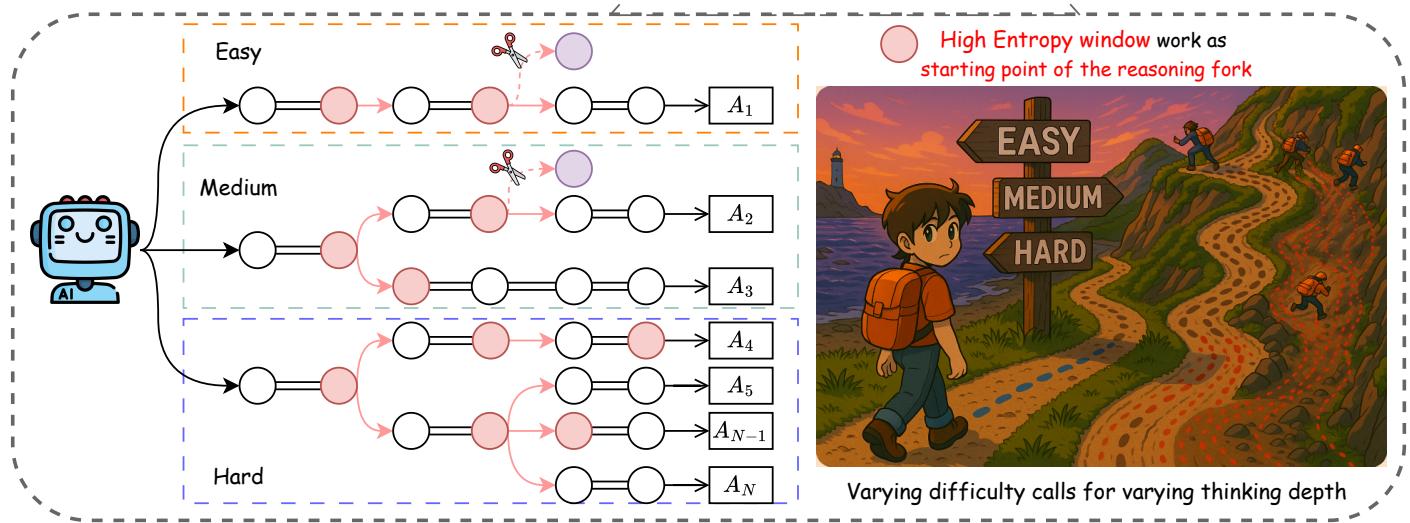
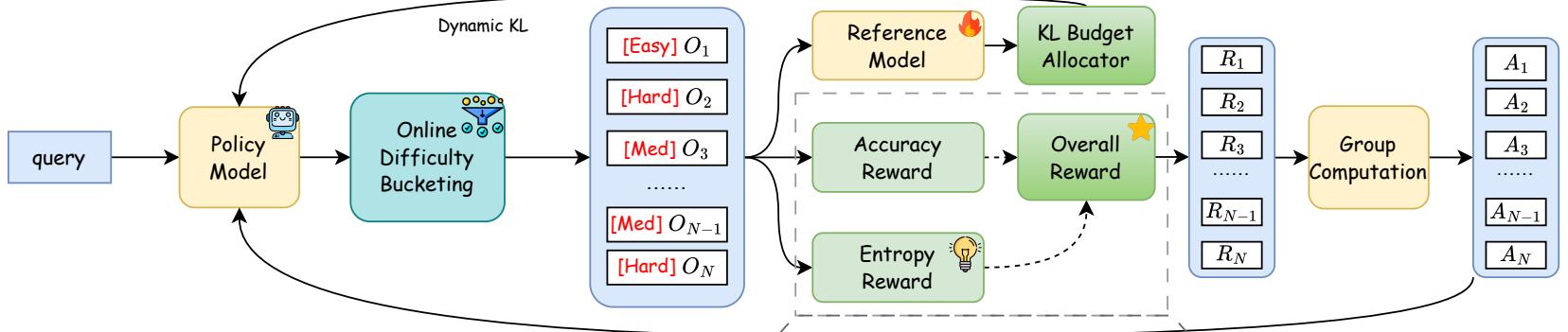


Training Stage 1: Adaptive Coldstart Fine-Tuning



Training Stage 2: Adaptive Entropy Policy Optimization(AEPO)