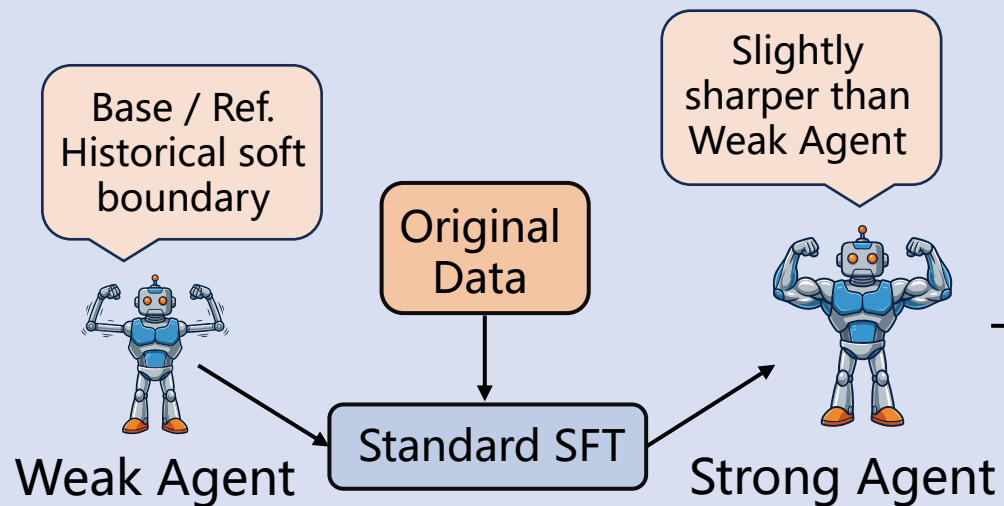
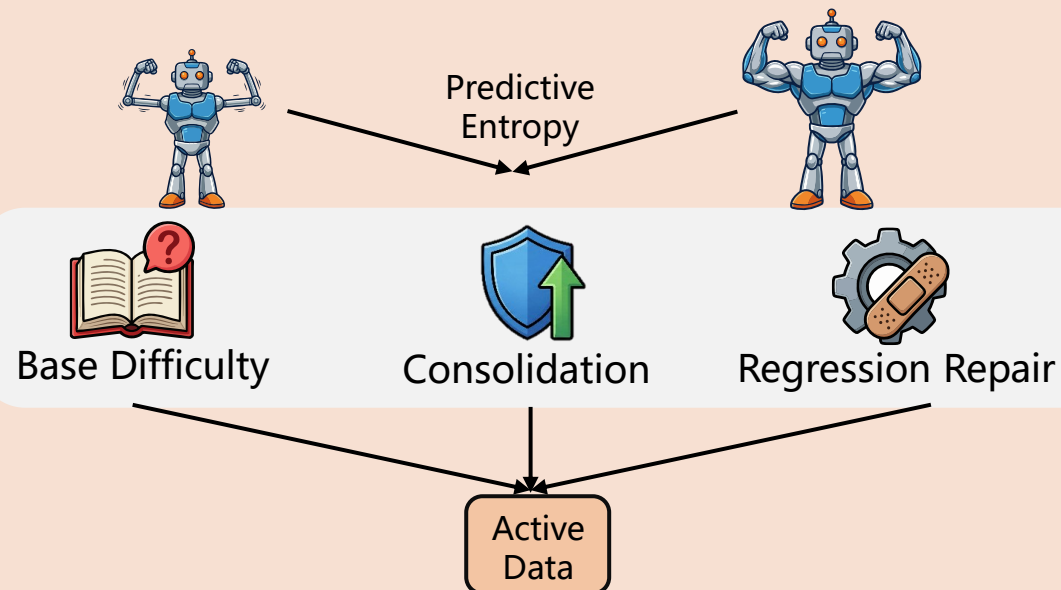


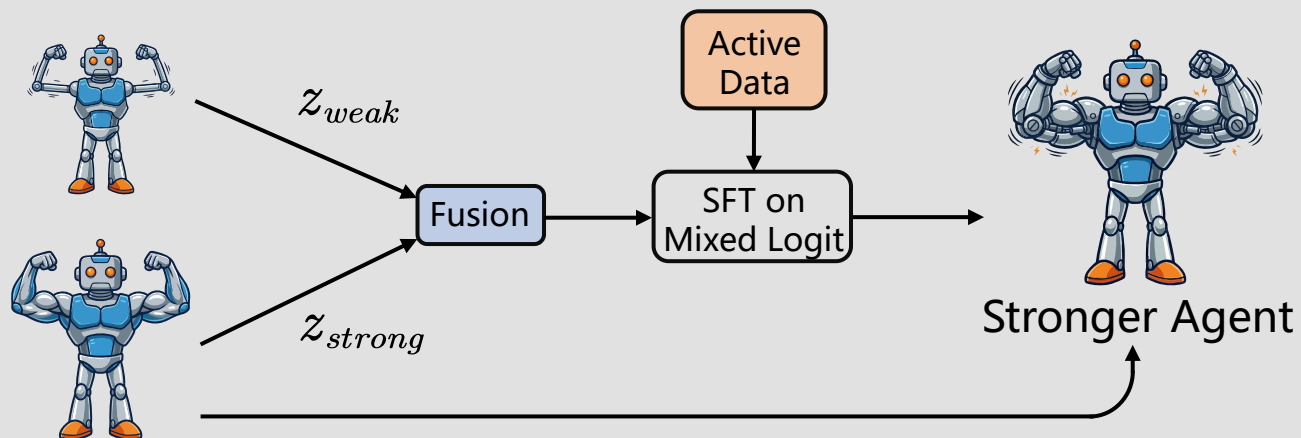
Stage 1 (Initialization Weak & Strong Agent)



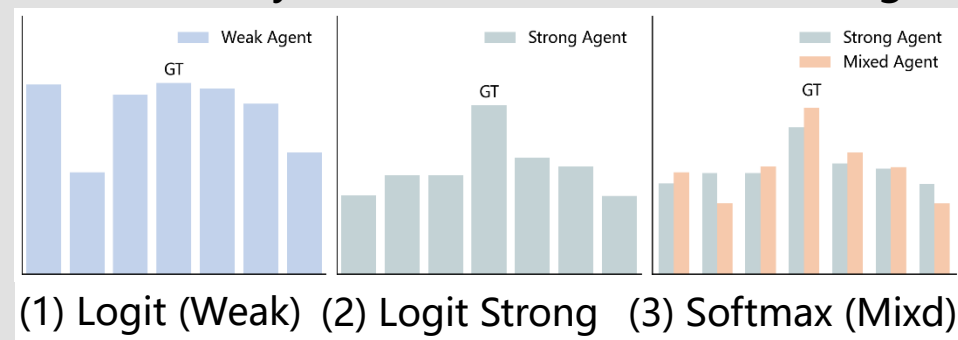
Stage 2 (Curriculum-Enhanced Data Activation)



Stage 3 (Joint Training via Logit Mixing, SFT)



Key Visualization of Joint Training



Mixing lifts hard negative mass while keeping GT highest; Strong Agent is only slightly sharper than Weak Agent.