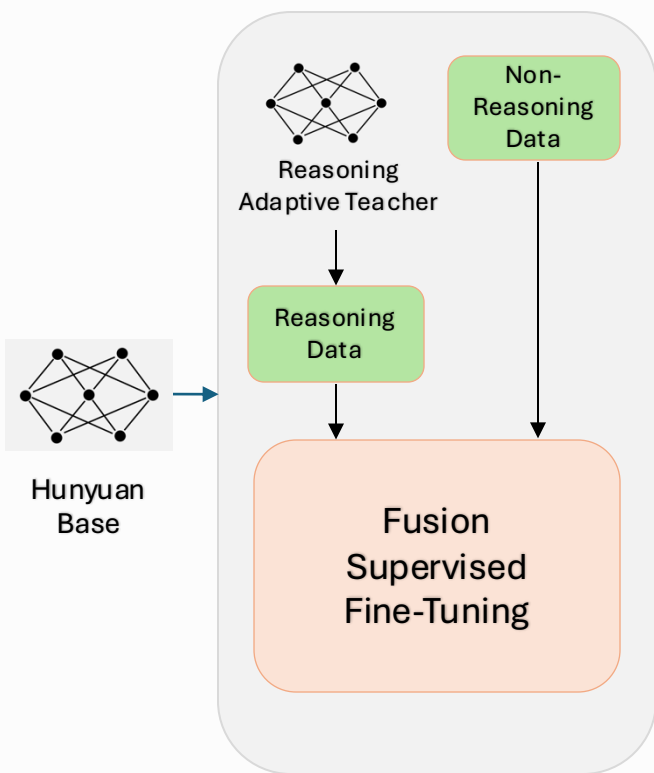
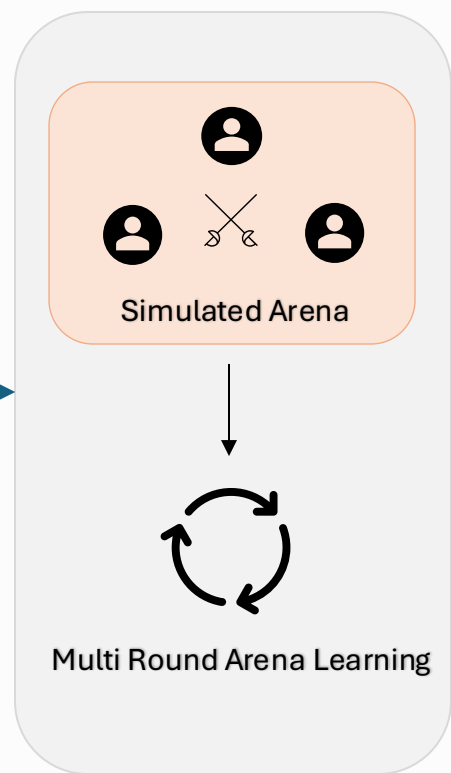


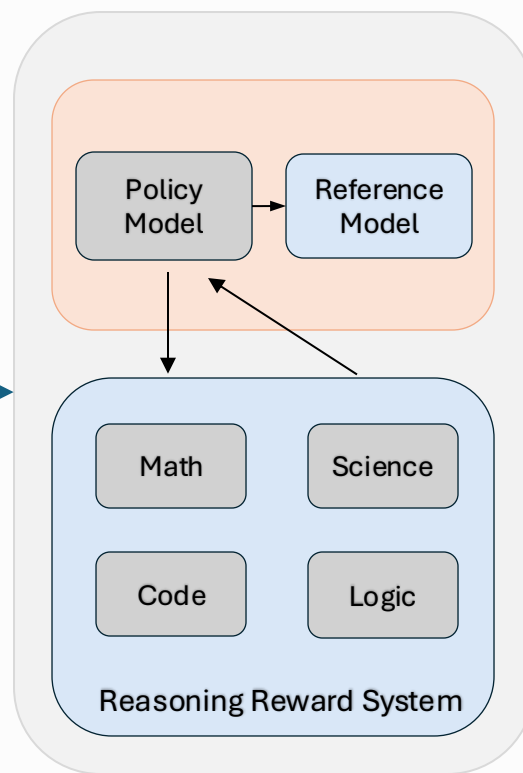
Step 1: Fusion Supervised Fine-Tuning



Step 2: Arena Learning SFT



Step 3: Reasoning GRPO



Step 4: General GRPO

