

(a) Standard Online RL



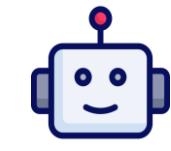
Trajectory



Outcome Reward

*Rare Success* !

(b) Experience Replay



Trajectory

Online

Buffer *Old Trajectory*

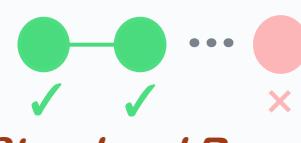
Outcome Reward

*Cross-Task* ✗

(c) Dense Reward



Trajectory



*Step-level Reward*

Dense Reward

*Cross-Task* ✗

(d) Evolving Memory (Ours)

Experience Memory

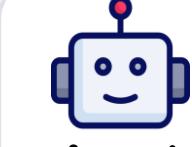
Successful Trajectory

Success Plan

Failure Pattern

Failed Trajectory

Guidance



Guided Trajectory

Guidance-aware Reward

*Cross-Task* ✓