

CliffWalking

Episode reward

0

-100

-200

-300

-400

-500

0

25

50

75

100

125

150

175

200

of episodes

- LORO. Cum. reward=-9213
- Online RL. Cum. reward=-37151
- Qwen-7B. Cum. reward=-101933
- Qwen-32B. Cum. reward=-1545806
- Random. Cum. reward=-404919

