

Reward as a function of time. $R_{\text{tot_naive}} = -1.1$ and $R_{\text{tot_smart}} = -1.92$

