

— Invalid action masking — Naive invalid action masking — Invalid action penalty, $r_{\text{invalid}} = -0.01$ — Invalid action penalty, $r_{\text{invalid}} = -1$

— Invalid action penalty, $r_{\text{invalid}} = 0$ — Invalid action penalty, $r_{\text{invalid}} = -0.1$