$r(s, a)$ (immediate reward)