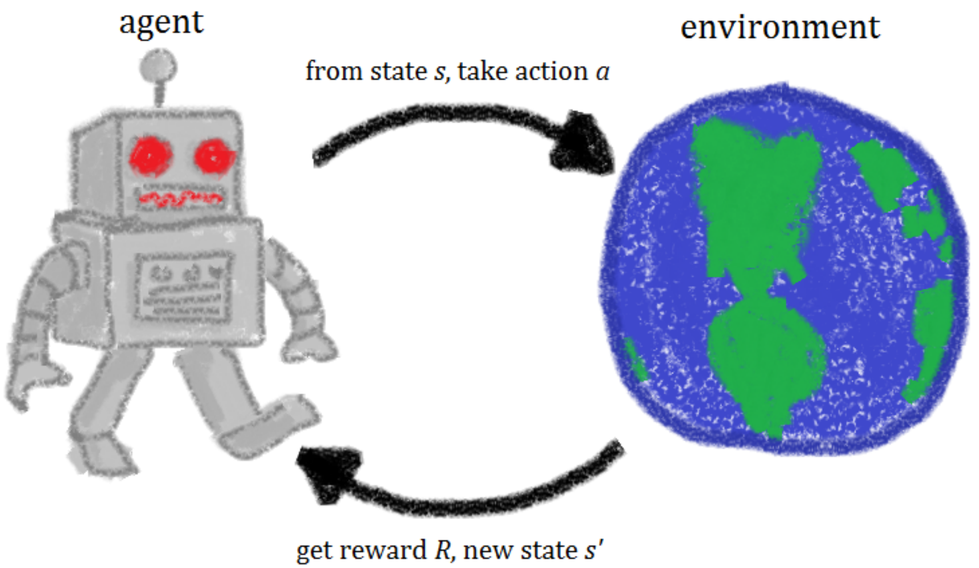
# RLHF




$$V(s) = \max_a \left( R(s,a) + \gamma \sum s' P\left(s,a,s'\right) V\left(s'\right) \right)$$