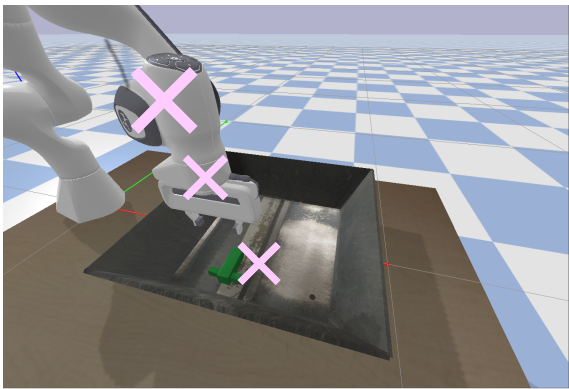
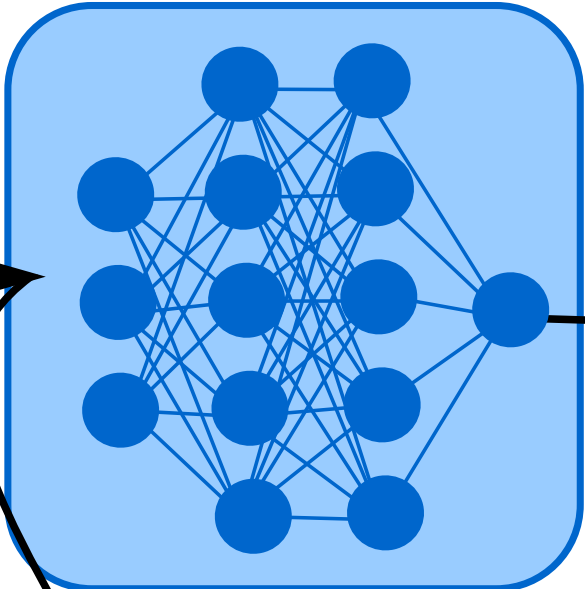


The environment exposes:
End effector world position
Object world position
Joint angles

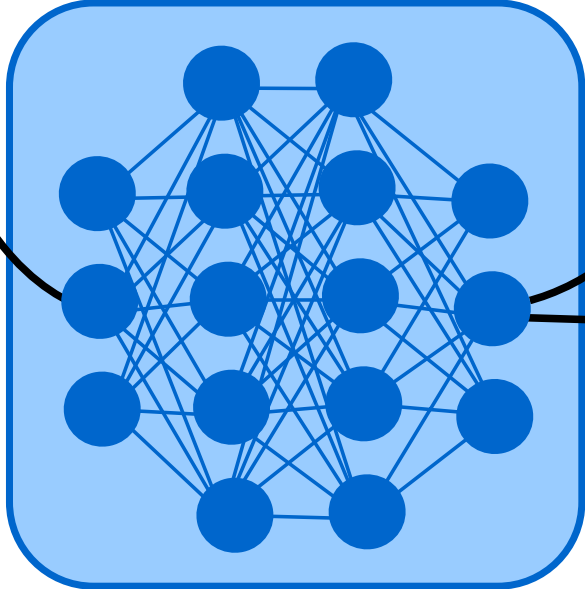


Observations
(joint and object states)
(6) Distance from object to each finger in world space

(3) Distance from object to end effector in world space
(1) Object Z coordinate
(3) End effector position in world space
(9) Joint angles of controlled joints



Q function
Single output
Q value
 $Q(s,a)$



Policy
Actions (joint control inputs)
9 degrees of freedom
Desired joint position controls