

Reward  $r$  (in the training phase)

