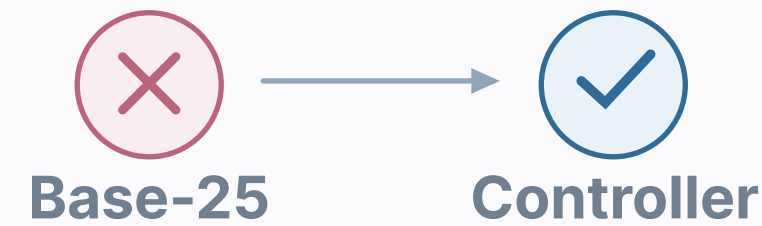


(a) Recovering policy failuresAmong episodes where Base-25 **failed****Weak set**

33 failures

5/33 (15.2%)

6/33 (18.2%)**Strong set**

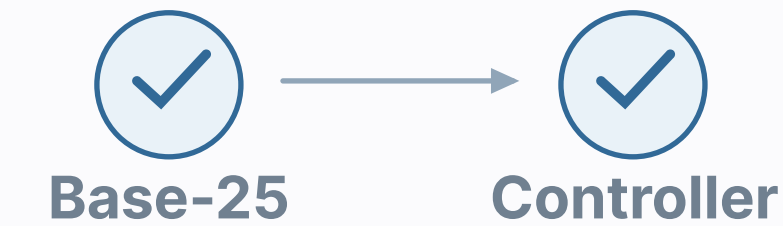
4 failures

2/4 (50.0%)

2/4 (50.0%)

0 25 50 75 100

Policy failures recovered (%)

(b) Preserving policy successesAmong episodes where Base-25 **succeeded****Weak set**

2 successes

1/2 (50.0%)

2/2 (100.0%)**Strong set**

11 successes

1/11 (9.1%)

9/11 (81.8%)

0 25 50 75 100

Policy successes preserved (%)