

eval successes

0.8
0.6
0.4
0.2
0.0

0.0

0.5

1.0

1.5

2.0

2.5

3.0

3.5

4.0

1e5

Steps

- SM (Ours)
- QL-SM (Ours)
- CRM
- CRM-RS
- HRM
- HRM-RS
- QL
- QL-RS