

平均回报

-4
-6
-8
-10
-12
-14

0

1

2

3

4

5

训练步数

1e6

Polar action
Descartes action

