

Final reward

200
175
150
125
100
75
50
25
0

Max reward

Exploration
phase

Exploitation
phase

Episodes

0

250

500

750

1000

1250

1500

1750

2000

