A
10 m
Target state (2 points)
Reward zone (1 point)
20 m
B
Centrality
0.15
0.1
0.05
0
1 2 3 4 5
Arena
C
s
s′
R(s,a)
a
−√(3)/3
−1
−2√(3)/3
D
P(outcome)
0.8
0.4
0
2points
1point
% Points
90
60
Stopping dist from goal (m)
3
2
1
0
Unrewarded
Rewarded
0 3 6
Arena complexity
E
P(2 points)
P(1 point)
% Points
P(rewarded)
1
0.5
0
0.6
0.4
0.2
100
80
60
40
0.8
0.6
0.4
0 3 6
Arena complexity
F
Stopping dist from goal (m)
Relative effect
0.4
0.2
0
-0.2
Turns
Length
Bearing