policy	reward	$(\pm std)$
B1	99.1	(± 129.6)
B2	90.9	(± 142.2)
B3	65.5	(± 137.3)
B4	176.0	(± 154.1)
B5	95.9	(± 144.9)
B6	168.8	(± 165.3)
B7	229.3	(± 157.1)
RL	310.8	(± 136.1)