

Mean episode reward

- cramped_room
- asymmetric_advantages
- coordination_ring
- soup_coordination

Environment steps 1e6

