benchmark rewards – Ø-Reward vs Gegner single_only max combo 2.8 random2 macro reward 2.6 1.8 1.6 20000 40000 60000 80000 100000 **Episode**