

TRPO static self-learned Hanabi policy

mean=1.1, stddev=0.3

