

TRPO static self-learned mini Hanabi policy

mean=2.0, stddev=1.483

