Craigslist

Ultimatum

Deal or No Deal

CaSiNo

Job Interview

Alliance Negotiation

Expected Return $J(\pi)$

Round

IterPO

ST-GRPO

MT-SFT

MT-KTO

MT-PPO

ARCHER

ARCHER-LC