Env - Red Bus/Blue Bus Paradox - Average Reward EXP3 - N 2 NEW - N 2 NEXP3 - N 2 0.95 → EXP3 - N 5 -**├-** NEW - N 5 -
NEXP3 - N 5 — EXP3 - N 10 — NEW - N 10 **─** NEXP3 - N 10 EXP3 - N 50 0.90 −>-- NEW - N 50 NEXP3 - N 50 0.85 Average 0.75 0.80 0.70 0.65 200 400 600 800 Steps