## WildfireResourceManagement Train & Test Metrics: Difficulty: 6, Task: 0 Entropy Individual Resource Count Reward for Moving Resources to Neighbours Reward for Moving Resources to Self Collective Performance Individual Performance 200 150 source source source source source - 2.4 Value Value training training training Value 10 -Value 8 100 · 100 training 0.7 test 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 Step Step Step Step Step Step **Episode Length** Curiosity Value Estimate Extrinsic Value Estimate **Cumulative Reward Curiosity Reward Extrinsic Reward** source source source source source 0.1 Value training training o.1 test training 0.0 -100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 Step Step Step Step Step Step Beta **Curiosity Forward Loss** Policy Loss Value Loss Learning Rate **Epsilon** 0.010 source source source source 0.4 2000.0 \frac{\text{qi ne}}{\text{N}} training training training training 0.005 0.15 = 100 -test test test — test 0.0000 0.10 0.000 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 100000 200000 300000 400000 Step Step Step Step Step Step **Curiosity Inverse Loss** source

100000 200000 300000 400000