WildfireResourceManagement Train & Test Metrics: Difficulty: 1, Task: 2 Entropy Individual Resource Count Reward for Moving Resources to Neighbours Reward for Moving Resources to Self Collective Performance Individual Performance 150 · source source source source 9 100 · training 0.8 training training - 00.4 Value 20 source training test 3.75 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 **Extrinsic Reward Episode Length** Curiosity Value Estimate Extrinsic Value Estimate **Cumulative Reward Curiosity Reward** source source source source source 0.1 2000 -9 1000 ω 2000 training training training training training 0.1 Value - 002 = 1000 ≤ 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 **Epsilon** Beta **Curiosity Forward Loss** Policy Loss Value Loss Learning Rate 0.010 -0.20 -20000 source source υ 0.0002 training 0.005 0.2 source source source 10000 -0.15 test training training training test 0.0000 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 - 00.4 3.75

100000 200000 300000 400000