WildfireResourceManagement Train & Test Metrics: Difficulty: 6, Task: 1 Entropy Individual Resource Count Reward for Moving Resources to Neighbours Reward for Moving Resources to Self Collective Performance Individual Performance source source source source 100 source 10 -400 -200 training Value training training training - 00.4 source training 0.7 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 **Episode Length** Curiosity Value Estimate Extrinsic Value Estimate Cumulative Reward **Curiosity Reward Extrinsic Reward** source source source source source training training 0.1 e test 50 training 0.0 test 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 **Curiosity Forward Loss** Policy Loss Value Loss Learning Rate **Epsilon** Beta 0.010 -0.20 -1000 source source source source aka ia ka 0.4 0.0002 Najre training training training training 0.005 0.15 500 test test test — test 0.0000 0.10 0.000 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 **Curiosity Inverse Loss** source 4.25 -9 4.00 -3.75 training

100000 200000 300000 400000