FDR for a1false-a2false-b1false-c1true, reward function 1 0.95 0.90 0.85 FDR bandit-opt 0.80 bandit-eps random 0.75 qlearning bayes-low bayes-high 0.70 mlfq tableau 0.65 ann 0 50 100 150 200 250 300 350 400

experiments run