FDR for a1false-a2false-b1false-c1true, reward function 1 0.95 0.90 0.85 0.80 bandit-opt bandit-eps 0.75 random qlearning bayes-low 0.70 bayes-high mlfq 0.65 tableau ann 0.60 50 200 0 100 150 250 300 350 400 experiments run