FDR for a1false-a2true-b1false-c1false, reward function 1 0.95 0.90 0.85 FDR bandit-opt 0.80 bandit-eps random qlearning 0.75 bayes-low bayes-high 0.70 mlfq tableau

0.65

0

50

100

150

200

experiments run

250

300

ann

400

350