

UCB1: best vs actual reward

