

ATR vs. Q-learning Iteration, $\alpha=0.8 \cdot \exp(-i)$

