

$$\frac{1}{N} \sum_{\tau \sim P_{\pi, T}} \left[R(\tau) \nabla \log P_{\pi, T}(\tau) \right]$$