Methods

1. Importance sampling

Drawback:

While importance sampling provides an unbiased estimate of the true objective, it often yields importance ratios with large and unbounded variance (Robert & Casella, 2004). Truncating or adjusting the importance weights can reduce the variance but introduces bias.

1. Multi-Agent Fingerprints

Use the training iteration number e and the rate of exploration as input to Q to solve nonstationarity.

My review:

Actions of others as inputs are more good.