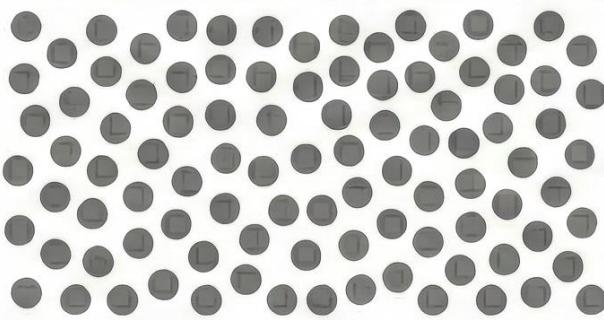


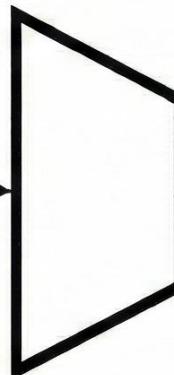
关键性筛选器 (Filter)

原始经验池
(Raw Experience Buffer)



$$C(s) = w_1 \mathcal{N}(x) + w_2 |\nu|$$

Data Stream

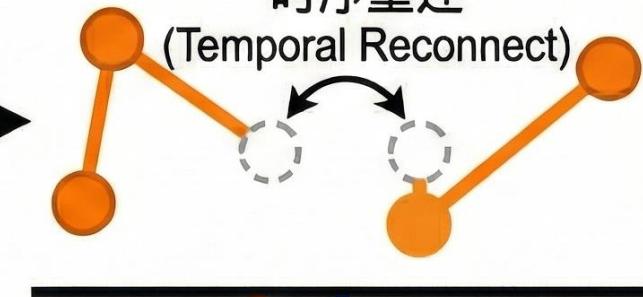


$$C(s) \geq \delta ?$$

YES

核心集
(Coreset)

时序重连
(Temporal Reconnect)



产生新数据
(Generate New Data)

环境交互
(Environment Interaction)

Target y
 $(r + \gamma \max Q')$

Loss Function
(MSE)

Deep Q-Network Optimization

Q-Values

Q_{left}
 Q_{wait}
 Q_{right}

Backprop

FC Layers + ReLU

采样
(Sampling)

更新后的策略
(Updated Policy π_θ)