

제목 : Deep Reinforcement Learning with Double Q-learning

저자 : Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, Daan Wierstra

1. 논문 주제

논문은 연속적인 행동 공간에서 강화 학습을 탐구

Deep Q-Learning의 성공적인 아이디어를 활용하여 연속적인 행동 도메인에서의 문제 해결을 시도.

이를 위해 결정론적 정책 기울기를 기반으로 한 액터-크리틱(actor-critic) 모델을 제시하였고, 이 모델이 연속적인 행동 공간에서 작동 가능함을 입증

동일한 학습 알고리즘, 네트워크 아키텍처 및 하이퍼파라미터를 사용하여 이 알고리즘이 20개 이상의 시뮬레이션 물리 과제를 안정적으로 해결하는 데 성공.

이 과제에는 고전적인 문제뿐만 아니라, cartpole swing-up, dexterous manipulation, legged locomotion, car driving 등이 포함

2. 주요 내용

2-1.액터-크리틱 모델의 제안:

결정론적 정책 기울기를 활용하여 연속적인 행동 공간에서의 액터-크리틱 모델을 제안

2-2.다양한 문제 해결:

동일한 알고리즘과 아키텍처를 사용하여 20개 이상의 다양한 시뮬레이션 물리 과제를 안정적으로 해결

2-3.Raw pixel inputs에서의 정책 학습:

많은 과제에서 raw pixel inputs에서 직접 정책을 학습할 수 있음을 입증