

[HOME](#)

3. DDPG

[Continuous control with deep reinforcement learning]

Timothy P. Lillicrap*, Jonathan J. Hunt*, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver & Daan Wierstra (2016)

[paper link](#)

[Abstract]

- Deep Q-Learning에서의 성공의 기반이 되었던 주요 아이디어들을 채용해 **continuous action domain**으로 옮겼다.
- deterministic policy gradient기반의 actor-critic, model-free 알고리즘이 continuous action spaces에서도 잘 돌아간다는 것 보인다.
- 동일한 학습 알고리즘, network 구조, hyper-parameter를 이용해 20가지가 넘는 물리 시뮬레이션 task를 모두 풀어냈다.
- 이 알고리즘은 이제 dynamics기반의 planning 알고리즘과 비교해 충분히 경쟁적인 성과를 내는 policy를 찾아낸다.

[Introduction]

AI분야의 본래 목적 중 하나는 바로 **가공되지 않은 고차원의 센서 데이터를 이용해 복잡한 문제를 푸는 것**이라고 할 수 있다.

최근에는 센서데이터 처리를 위해 발전된 딥러닝 기술을 사용하기 시작했으며, 이것이 강화학습과 결합되어 Deep Q-Network가 탄생하게 되었다.

주목할 점은 **action-value function**을 추정하기 위해 **deep neural network function approximator**를 사용한다는 것이다.

그러나 DQN은 high-dimensional observation spaces를 이용해 문제를 풀지만, **discrete하고 low-dimensional action spaces만을 다룰 수 있었다.**

DQN은 action-value function을 최대화 하는 action을 찾는 방식에 의존하기 때문에 **continuous domain에는 바로 적용할 수 없다.**

continuous valued case에서는 action-value function의 최대화 과정에서 매 스텝마다 최적화 프로세스를 요구하기 때문이다.

때문에 DQN에서는 continuous domain문제에 적용하기 위한 방법으로 action space를 discrete한 action으로 쪼개어 적용하였다.

그러나 이러한 방법은 많은 **한계점(limitations)**을 지닌다. 바로 **차원의 저주(the curse of dimensionality)** 때문이다.

[Background]

[Algorithm]

[Results]

[Related Work]

[Conclusion]