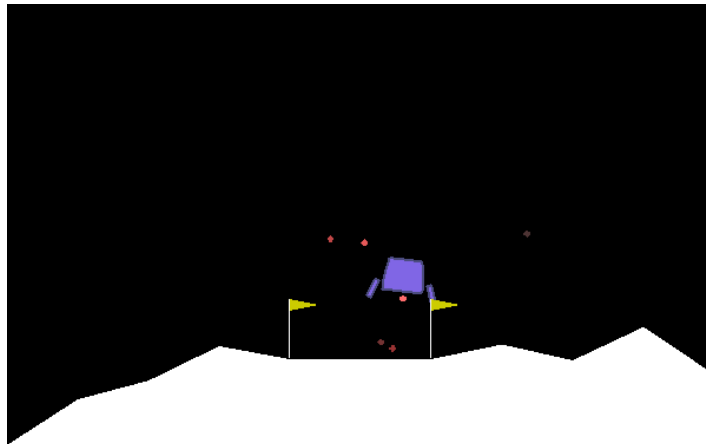


달 착륙(Lunar Lander) Competition



1. 상태(state)

[
위치 x좌표,
위치 y좌표,

속도 x성분,
속도 Y성분,

기울어진 각도,
각속도,

왼쪽 다리가 땅에 닿았는지 여부,
오른쪽 다리가 땅에 닿았는지 여부
]
로 총 8차원 벡터가 주어집니다.

2. 행동(action)

- 0 - 아무 것도 하지 않음.
- 1 - 왼쪽 엔진을 점화 (왼쪽으로 가는 거 아님)
- 2 - 메인 엔진을 점화 (아래쪽으로 가는 거 아님)
- 3 - 오른쪽 엔진을 점화 (오른쪽으로 가는 거 아님)

3. 보상(reward)

- 각 다리가 땅에 닿으면 +10
- 메인 엔진이 점화된 상태면 각 프레임당 -0.3
- 보조 엔진이 점화된 상태면 각 프레임당 -0.03
- 착륙에 성공하면 +200
- 에피소드가 끝날 때 땅에 충돌하면 -100
- 에피소드가 끝날 때 땅에 충돌하지 않았다면 +100

4. 에이전트에 영향을 미치는 요소들

- 학습률(Learning rate)
학습률은 보통 0에서 1 사이의 값을 정합니다. 너무 큰 값으로 설정하면 학습이 되지 않고 발산할 수 있으며, 너무 작은 값으로 설정하면 학습이 매우 느려질 수 있습니다. 따라서 “적절한” 값을 찾는 것이 중요합니다.
- Optimizer
optimizer엔 SGD, RMSprop, Adagrad, Adadelata, Adam, Nadam 등이 있습니다. Optimizer의 차이는 학습 속도 혹은 학습 성능의 차이를 야기합니다. 따라서 “적절한” optimizer를 찾는 것이 중요합니다.

Tip: optimizer는 <https://keras.io/optimizers/>를 참고하세요.

- epsilon decay
학습함에 따라 epsilon 값을 적절히 줄이는 것이 중요합니다.