

Concept

대부분의 사람들의 경우 이상적인 라인을 그려서 DeepRacer로 하여금 그 선과 최대한 가깝게 주행할 때 더 큰 Reward를 줌으로써 실물을 학습시키는 방법을 사용한다.

우리는 이런 방법 말고 다른 방법을 사용해볼 것이다. 차량을 중심으로 가상의 큰 원을 그리도록 하겠다. 그리고 이 원을 기하학적 방법을 사용해서 강화학습을 할 것이다. 기하학적이라 해봤자 피타고라스의 정리, 극좌표, xy좌표에 대한 개념정도만 나오고 있다. 차량을 중심으로 원을 하나 그려서 그 원과 중앙성(waypoint)가 만나는 지점을 찾는다. 그리고 그 지점을 이용해서 우리가 조향해야 할 올바른 방향을 구하고 그 방향으로 차량이 가깝게 향할수록 더 큰 Reward를 받을 수 있도록 하겠다.

여기서 특이한 점은 discount factor를 0.5를 주었다는 점이다. 다른 hyperparameter는 기본 값에서 딱히 변경하지 않았다. (참고로 discount factor의 기본값은 0.999이다, 이 Reward Function의 제작자는 discount factor를 더 줄일 수도 있을 것이라고 설명하고 있다. 아마도, 그렇게 하면 학습의 속도가 더 빨라지지 않을까 싶다.)

이 학습법의 가장 큰 장점은 빠른 학습이 가능하다는 점이다. 학습의 속도는 즉각적인 Reward를 받을 수 있을 때 더욱 빨라질 수 있다. 만약 에피소드가 하나 끝날 때까지 기다린 후 그 에피소드에 대한 Reward를 부여하는 경우, 즉각적인 Reward를 받을 수 없기 때문에 학습의 속도가 느려진다. 반면, 조향의 방향과 waypoint를 이용한 function은 오롯이 카메라를 통해 얻은 이미지를 사용해서 얻게 되는 Reward Function이다.

즉, 바로바로 Reward를 부여할 수 있기 때문에 학습의 속도가 빠를 것이다. Cutting Corner라는 말이 나오는데, 이 말은 무슨 말인지는 잘 모르겠다. 아마, 코너에서는 속도를 줄인다거나 코너는 대충처리해라 이런 뜻같은데, 코너에서 속도를 줄여가 더 적합한 표현이 아닐까 생각한다. 코너에서는 차량을 중심으로 그려지는 원의 반경을 더욱 크게 만들라고 말하고 있는데, 이는 시뮬레이션을 통해서 알아가야 될 듯 싶다.

글로만 봐서는 무슨 말을 하고 싶은 건지 잘 모르겠다.

링크 :

AWS Deepracer-How to train a model in 15 minutes

In a previous article I described how I got into DeepRacer and what it is. This article is about the technical parts: How the scoring function was designed and why it works. Almost all

M <https://medium.com/@falktandetzky/aws-deepracer-how-to-train-a-model-in-15-minutes-a07ab77fb793>

