**인공지능개론 14주차 과제**

-벨만 기대 방정식(Bellman Expextation Equation)-

20211848 안주형

* 주제 선정 이유

강화학습에 대한 강의 자료를 보았는데, 강의를 들을 때부터 관심이 가던 선택할 때 가치를 더 높이는 쪽으로 높은 쪽으로 구할 때 그 가치를 구하는 방법이 궁금했다. 그리고 영상에서 가치를 구하는 작업을 벨만 기대 방정식을 이용하여 정리하는 것을 보고 주제로 삼게 되었다.

* 벨만 기대 방정식이란?

벨만 기대 방정식은 에이전트가 최적의 정책을 찾는 과정에서 각 상태의 가치를 계산하는데 사용된다. 벨만 방정식은 동적 계획법의 기본 원칙을 따르며, 주어진 상태에서 미래의 보상을 예측하는 데 사용된다. 강의 영상 중 교수님께서 서양사람들은 쉬운 걸 어렵게 쓰는 걸 잘한다고 하셨는데 그 말이 백번 공감이 갔다.

폰트, 화이트, 텍스트, 서예이(가) 표시된 사진

자동 생성된 설명

위 공식이 벨만 기대 방정식인데 의미하는 문자가 너무 많아서 상당히 이해하기 어려웠고, 완벽하게 이해하기까지는 여러 번 반복하는 시간이 필요할 것 같다.

이해한 걸로 보았을 때 Vn(s)가 현재 방의 가치이고 Vn(s’)가 다음 방의 가치를 예상 것이다. 여기서 멀리서 구해온 것에 대한 감가율과 방향의 확률 그리고 액션을 취해서 얻은 값 등등을 고려하여 다음 방의 가치를 구한다고 생각된다.

검색을 통해 알아 보았을 때 Vn(s) 말고도 행동 가치 함수 Q(s,a)도 따로 있는 것 같고 원래의 함수를 벨만 방정식으로 적용하는 듯하다 마찬가지로 상태 가치 함수를 벨만 기대 방정식에 적용한 것이 위의 방정식과 같은 모습이다. 이 방정식들은 동적 계획법과 강화 학습 알고리즘에서 핵심적인 역할을 한다.