|  |  |
| --- | --- |
| 교육 제목 | 강화 학습 |
| 교육 일시 | 2021.11.29 |
| 교육 장소 | 영우글로벌러닝 |
| **교육 내용** | |
| 오전 | 에이전트는 간단히 환경이라는 뜻을 담고 있다  에이전트의 알고리즘 - 라이브러리  에이전트  policy은 행동과 상태를 수집하여 어디로 가야할지 정한다  value function은 취한 행동이 대하여 좋은지 나쁜지 나열한다  model은 에이전트가 이해하고 있는 환경상태를 나타낸다  엑터 크레틱은 policy과 value function이 모두 존재한다  MDP는 이해를 한 환경을 전제조건으로 둔다  모든 상황을 포괄적으로 담고 있다는 개념이다 |
| 오후 | process는 s,p로 나뉜다  Decision은 S,A,P,R으로 나뉘며 포함을 시킨다  Decision의 경우  내가 이러한 선택을 했을 때 기대할 수 있는 값은 무엇인지?  내가 이러한 행동을 했을 때 기대할 수 있는 값은 무엇인지?  식으로 도출해낸다 |