|  |  |
| --- | --- |
| 교육 제목 | AI 프로젝트 기반 S/W 전문가 양성과정 |
| 교육 일시 | 21.11.29 |
| 교육 장소 | YGL |
| **교육 내용** | |
| 오전 | 1. 기계 학습 마지막날 2. 강화 학습 3. Markov decision process   Fully observable environments  Partially observable environments = 부분적 100% 알지 못하기 때문에 확률 개념 등장   1. Agent = policy : agent’s behavior function Value function : how good is each state and or action model : agent’s representation of the environment (강화학습의 모델은 agent가 이해하고 있는 환경) 2. Exploitation : 개척한 곳만 가는 느낌 Exploration : 개척하는 것 3. policy = 상태와 행동들을 모아 놓은 것 stochastic policy : 확률 4. value function = future reward를 찾기 위한 계산 5. Model = predict 예상 / Estimation 예측 6. Reinforcement learning is like trial-and-error learning 7. Markov Decision Processes Markov Processes = Markov chain tuple <S, P> S = a finite set of states P = a state transition probability matrix |
| 오후 | 1. fully observation에서 2. bellman 방정식   v(s) = 앞으로 일어날 일들의 집합  policy =   1. expectation 방정식 2. optimal value 함수 3. Dynamic sequential or temporal component to the problem  Programming optimising a “program”, i.e. a policy c.f. linear programming |