강화학습 모르겠는거

1. MDP는 마르코프 속성을 전재로 하는데, 즉 미래의 일은 직전(현재)만이 영향을 준다. 다시 말해 s다음 s’는 s가 되기까지 과정과 전혀 상관없다. 그렇다면 sc2라는 환경에서 마르코프 속성이 없을텐데(초반 일꾼 잡힌거랑 후반 일꾼 잡힌거랑 다르듯이), 알파스타는 어떤 원리로 구현 되었는가(MDP가 적용이 안된 강화학습인가? 하지만 대부분의 강화학습은 MDP의 일부라고 알고 있는데..)
2. 그럼 MDP는 정말 특별한 상황에서의 이론적 강화학습인건가? 마르코프 속성 자체가 특별한 상황만 해당하는 것 같은데. + MDP는 그럼 모든 환경의 전이 확률이나 보상 같은걸 다 아는 한정된 상황에서의 강화학습인가?