q-learning

강화학습의 일종으로 q값(보상)을 설정해서 목표까지의 보상이 가장 큰 최적화된 경로를 찾는 방법으로 status(현재 상황 q값이라고 봐도 됨) action(행동) Episode(시도 횟수), step(action을 몇번 하는가) 등의 개념을 사용해 구현할 수 있다. 또한 완전히 한 경로에 최적화 되는 문제를 해결하기 위해 Exploration(새로운 경로 탐색)이라는 개념을 도입해 일정 확률로 랜덤의 새경로를 찾아가는 방식을 사용하며 이때는 e-greedy 방식을 이용해 수행한다. 또한 action마다 보상이 주어지며 할인률을 통해 미래보상을 고려합니다.

e-greedy

epsilon 이라는 초매개변수를 통해 확률값을 조정하여 1-epsilon의 확률로 기댓값이 가장 큰 것으로 정하고 epsilon의 확률로 랜덤으로 다른 것을 선택하는 알고리즘

여기서 제안하고자 하는 프로젝트는 frozen lake 게임으로 장애물을 패해서 목적지에 도달해야 하는 게임으로 강화 학습을 통해서 최적화된 경로를 통해 목적지에 도달하도록 학습시키고자 하는 것입니다. 또한 시각적으로 각 action을 보여준다면 발표를 하거나 할 경우 보다 더 명확하게 보여줄 수 있다는 장점 또한 갖고 있어서 frozen lake 게임을 구현해보는 것을 제안합니다.