**수치해석**

**팀 Project**

**김수빈**

**xxx**

**xxx**

**xxx**

\*matlab 코드에 대한 내용 이해 및 분석 🡪ppt에 넣을 내용

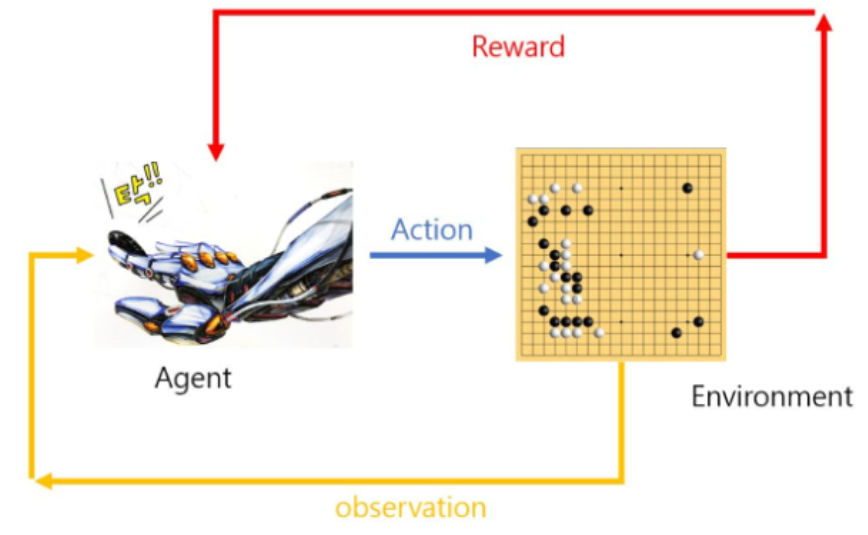
(1)Q-Learning

-Model 없이 학습하는 강화학습 알고리즘

-유한한 마르코프 결정 과정(FMDP)에서 Agent가 특정 상황에서 특정 행동을 하라는 최적의 수단을 배우는 것을 목표로 함

-현재 상태로부터 시작하여 모든 연속적인 단계들을 거쳤을 때 전체 보상의 예측값을 극대화시킴

-이때, “Q”라는 단어는 현재 상태에서 취한 행동의 보상에 대한 quality를 상징

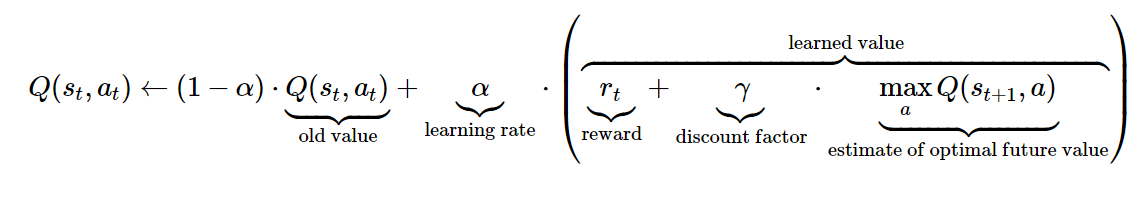


(2)Q-Learning Algorithm

알고리즘이 시작되기 전에 Q 함수는 고정된 임의의 값을 가짐.

매 시간마다 Agent는 행동 를 선택하게 되고, 보상 를 받으며 새로운 상태 로 전이하고, Q값이 갱신됨.

이전의 값과 새로운 정보의 가중합을 이용하는 Value Iteration Update 기법이 이 알고리즘의 핵심이라 할 수 있음.



%이 코드는 로봇이 왼쪽, 오른쪽, 위, 아래 방향으로만 이동하여 목적지에 도달해야하는 미로의 예를 사용하여 Q-Learning Algorithm을 보여주고 있습니다.

%각 단계에서 로봇 동작의 결과에 따라 학습이 잘 되었는지의 여부를 다시 학습하고 전체 프로세스가 목적지에 도달할 때까지 반복됩니다.

%이 시점에서 프로세스가 다시 시작되어 학습한 내용을 확인하고, 첫 번째 패스에서 수행한 불필요한 동작을 잊을 수 있습니다.

%크기가 작은 미로에서는 수렴 속도가 빠르며, 큰 미로에서는 시간이 오래 걸릴 수 있습니다.

%네 개의 m-파일이 있습니다.

%QLearning\_Maze\_Walk.m-: 선택한 미로에서 Q-Learning Algorithm의 작동을 보여줍니다.

%Random\_Maze\_Walk.m-: 비교를 위한 무작위 선택의 작동을 보여줍니다.

%Read\_Maze.m-: 입력으로 제공된 미로를 읽고 처리를 위해 숫자 표현으로 변환합니다.

%Textscanu.m-: 원시 미로 텍스트 파일을 읽습니다.

1. QLearning\_Maze\_Walk.m

