Assignment 3

2015004448 김정욱

# 코드설명

dx=[0,1,0,-1], dy=[1,0,-1,0]는 4가지 방향을 나타낸다.

arr은 미로를 입력받은 변수이고, Q는 이차원배열에 각각 딕셔너리로 방향4군데에 대한 Q(s,a)함수값을 가진다.

1. learning()

스크린샷이(가) 표시된 사진

자동 생성된 설명

시작점 좌표를 받아와서 H나 G에 도착할 때 까지 좌표를 이동하며 학습한다.

1. 홀이나 골에 도착하면 종료시킨다.
2. 4방향 중 랜덤으로 하나를 선택한다. (Q함수가 존재하는 갯수 중에서 랜덤)
3. 다음 방향으로 갈때 홀이나 골이면 리워드값 갱신
4. 아니면 Q함수값중 최대값\*0.5로 갱신
5. 다음 좌표로 Q-learning

2. FrozenLake()

텍스트, 스크린샷이(가) 표시된 사진

자동 생성된 설명

m, n, inp를 받아온다. 시작점을 찾아서 좌표를 sx, sy에 저장한다.

4방향중 범위를 넘어가지 않는 방향에 대해 Q값을 0 으로 초기화한다.(범위를 넘어서면 생성하지 않음)

시작점의 Q함수 값이 갱신된 이후로 10만번 반복하여 learning함수를 실행한다.

(시작점의 Q함수 값이 모두 바뀐다면 미로를 다 탐색했다고 볼 수 있다.)

시작점의 반환해준다.

3. play()

텍스트, 스크린샷이(가) 표시된 사진

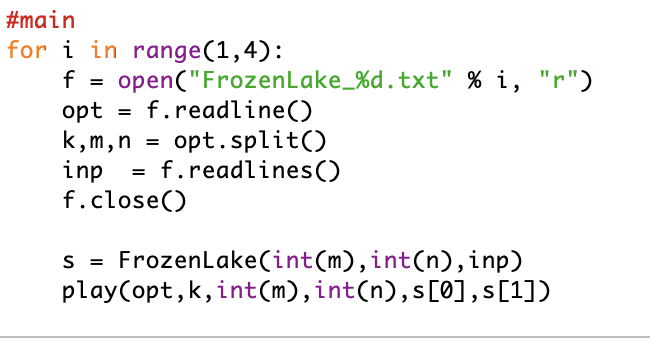
자동 생성된 설명

G지점에 도착할 때 까지 좌표를 이동하며 탐색하며 지나간 위치는 R로 바꿔준다.

현 위치에서 Q함수의 최대값을 찾아서 그 다음 위치를 결정한다.

딕셔너리 형태로 저장되어있기 때문에 인덱스 값을 받아오기 위해서 value리스트 key리스트를 만들어 사용.

4. 메인함수



입력을 받고 k,m,n,미로를 구분해준다. Q-learning을 위한 Frozenlake함수와 경로를 찾고 출력하기 위한 play함수를 실행시킨다.

# 결과

