

강화학습 구현과 실행

강화학습에는 여러 기법이 있는데 여기에서는 DQN(Deep Q-Network)를 사용하여 강화학습을 진행하였습니다

DQN은 기존의 Q-Learning에 신경망을 결합한 것입니다

Q-Learning은 주어진 상태에서 행동을 수행하면서 미래의 효율적인 기댓값을 예측하는 Q 함수를 학습하면서 최적의 정책을 학습하는 기법입니다

여기서는 DQN을 이용하여 떨어지는 블록을 바(bar)로 받아내는 간단한 블록 받기 게임을 학습시켜 에이전트에게 게임을 플레이 시킬 것입니다

설치

```
In [5]: !pip install tensorflow
!pip install matplotlib
```

```
zsh:1: command not found: pip
zsh:1: command not found: pip
```

import

```
In [34]: import tensorflow as tf
import numpy as np
import random
import math
import os
tf.compat.v1.disable_eager_execution()
# tensorflow v2에서 사용법인 바뀐 함수들은 앞에 tf.compat.v1.를 붙여서 사용해준다
```

파라미터 설정

```
In [35]: epsilon = 1 # 랜덤하게 행동할 확률
epsilonMinimumValue = 0.001 # epsilon의 최소값
nbActions = 3 # 행동의 개수 (왼쪽, 대기, 오른쪽)
epoch = 1001 # 게임 반복횟수
hiddenSize = 100 # 히든 레이어 뉴런 개수
maxMemory = 500 # 게임내용을 기억하는 최대 개수
batchSize = 50 # 학습시 데이터 묶음 개수
gridSize = 10 # 게임 화면 크기
nbStates = gridSize * gridSize # 상태 개수 (게임 화면 픽셀의 개수)
discount = 0.9 # 감소값
learningRate = 0.2 # 학습률
```

딥러닝 모델 설정

입력레이어는 nbStates(100)개 히든레이어는 hiddenSize(100)개 출력레이어는 nbAction(3)개를 가지는 딥러닝 모델을 만들어 준다
CNN을 사용한다

입력 레이어

X = tf.compat.v1.placeholder(tf.float32, [None, nbStates]) # 입력값

placeholder는 변수를 선언할 때 값을 바로 주는 것이 아닌 나중에 값을 던져줄 수 있도록 공간을 미리 만들어 주는 것이다

따라서 X는 데이터 유형이 float32이고 첫번째 차원의 수는 정해져있지 않고(가변적) 두번째 차원의 수는 nbStates의 값을 가지는 placeholder가 된다

W1 = tf.Variable(tf.random.truncated_normal([nbStates, hiddenSize], stddev=1.0 / math.sqrt(float(nbStates)))) # 가중치

입력 레이어의 가중치를 나타내는 W1을 생성해 준다

변수의 초기화는 랜덤값으로 주어지는데 이 랜덤값은 양쪽 끝이 잘려있는 정규분포에서 가져오게 된다 이 정규분포의 표준편차는 $\frac{1}{\sqrt{nbStates}}$ 가 된다

이러한 랜덤값들로 nbStates x hiddenSize의 크기를 가지는 행렬을 채워준다

b1 = tf.Variable(tf.random.truncated_normal([hiddenSize], stddev=0.01)) # 편향

활성화 난이도를 조절해주는 b1을 생성해준다

학습데이터가 가중치와 계산되어 나온 값에 더해주어 활성화 난이도를 조절해주는 역할을 해준다

여기서도 양쪽 끝이 잘려있는 정규분포를 이용하게 되는데 이 정규분포의 표준편차는 0.01이 된다

hiddenSize개의 랜덤값들을 만들어준다

input_layer = tf.nn.relu(tf.matmul(X, W1) + b1)

input layer의 출력값을 만들어준다

우선 X과 W1의 행렬곱을 한 다음 b1을 더해준다

그 다음 활성화 함수에 넣어주는데 여기서 활성화 함수는 ReLU 함수가 사용된다

ReLU 함수: $f(x) = \max(0, x)$

입력값인 X와 가중치인 W1을 행렬곱을 하게 되면 X는 None x nbStates의 크기를 가지는 행렬이고 W1은 nbStates x hiddenSize의 크기를 가지는 행렬이므로 결과는 None x hiddenSize의 크기를 가지는 행렬이 나오게 된다

이렇게 행렬곱을 하게되면 각각의 입력 레이어 노드에 들어가는 입력값들을 만들어낼 수 있다

이 값에 편향인 b1을 더해준 뒤 활성화 함수에 넣어주면 각각의 입력 레이어 노드의 출력값들을 만들어낼 수 있다

```
In [36]: X = tf.compat.v1.placeholder(tf.float32, [None, nbStates]) # 입력값
W1 = tf.Variable(tf.random.truncated_normal([nbStates, hiddenSize], stddev=1.0 / math.sqrt(float(nbStates)))) # 가중치
b1 = tf.Variable(tf.random.truncated_normal([hiddenSize], stddev=0.01)) # 편향
input_layer = tf.nn.relu(tf.matmul(X, W1) + b1) # 출력값
```

히든 레이어

W2 = tf.Variable(tf.random.truncated_normal([hiddenSize, hiddenSize], stddev=1.0 / math.sqrt(float(hiddenSize))))

히든 레이어의 가중치를 나타내는 W2를 생성해준다 입력 레이어와 마찬가지로 변수의 초기화는 랜덤값으로 주어지는데 이 랜덤값은 양쪽 끝이 잘려있는 정규분포에서 가져오게 된다

이 정규분포의 표준편차는 $\frac{1}{\sqrt{hiddenSize}}$ 가 된다

이러한 랜덤값들로 hiddenSize x hiddenSize의 크기를 가지는 행렬을 채워준다

b2 = tf.Variable(tf.random.truncated_normal([hiddenSize], stddev=0.01)) # 편향

활성화 난이도를 조절해주는 b2을 생성해준다

입력 레이어와 마찬가지로 학습데이터가 가중치와 계산되어 나온 값에 더해주어 활성화 난이도를 조절해주는 역할을 해준다

여기서도 양쪽 끝이 잘려있는 정규분포를 이용하게 되는데 이 정규분포의 표준편차는 0.01이 된다
hiddenSize개의 랜덤값들을 만들어준다

hidden_layer = tf.nn.relu(tf.matmul(input_layer, W2) + b2) # 출력값

hidden layer의 출력값을 만들어준다

우선 input_layer(입력 레이어의 출력값)과 W2의 행렬곱을 한 다음 b2을 더해준다

그 다음 활성화 함수에 넣어주는데 여기서도 활성화 함수는 ReLU 함수가 사용된다

ReLU 함수: $f(x) = \max(0, x)$

입력 레이어의 출력값인 input_layer와 가중치인 W2을 행렬곱을 하게 되면 input_layer는 None x hiddenSize 크기를 가지는 행렬이고 W2은 hiddenSize x hiddenSize의 크기를 가지는 행렬이므로 결과는 None x hiddenSize의 크기를 가지는 행렬이 나오게 된다

이렇게 행렬곱을 하게되면 각각의 히든 레이어 노드에 들어가는 입력값들을 만들어낼 수 있다

이 값에 편향인 b2을 더해준 뒤 활성화 함수에 넣어주면 각각의 히든 레이어 노드의 출력값들을 만들어낼 수 있다

```
In [37]: W2 = tf.Variable(tf.random.truncated_normal([hiddenSize, hiddenSize], stddev=1.0 / math.sqrt(float(hiddenSize)))) # 가중치
b2 = tf.Variable(tf.random.truncated_normal([hiddenSize], stddev=0.01)) # 편향
hidden_layer = tf.nn.relu(tf.matmul(input_layer, W2) + b2) # 출력값
```

출력 레이어

W3 = tf.Variable(tf.random.truncated_normal([hiddenSize, nbActions], stddev=1.0 / math.sqrt(float(hiddenSize))))

출력 레이어의 가중치를 나타내는 W3을 생성해준다

히든 레이어와 마찬가지로 변수의 초기화는 랜덤값으로 주어지는데 이 랜덤값은 양쪽 끝이 잘려있는 정규분포에서 가져오게 된다

이 정규분포의 표준편차는 $\frac{1}{\sqrt{hiddenSize}}$ 가 된다

이러한 랜덤값들로 hiddenSize x hiddenSize의 크기를 가지는 행렬을 채워준다

b3 = tf.Variable(tf.random.truncated_normal([nbActions], stddev=0.01))

활성화 난이도를 조절해주는 b3을 생성해준다

히든 레이어와 마찬가지로 학습데이터가 가중치와 계산되어 나온 값에 더해주어 활성화 난이도를 조절해주는 역할을 해준다

여기서도 양쪽 끝이 잘려있는 정규분포를 이용하게 되는데 이 정규분포의 표준편차는 0.01이 된다

bnActions개의 랜덤값들을 만들어준다

output_layer = tf.matmul(hidden_layer, W3) + b3

out layer의 출력값을 만들어준다 hidden_layer(히든레이어의 출력값)과 W3의 행렬곱을 한 다음 b3을 더해준다

회귀 문제이기 때문에 출력 레이어에서는 활성화 함수를 사용하지 않습니다 (항등함수라고도 함)

히든 레이어의 출력값인 hidden_layer와 가중치인 W3을 행렬곱을 하게 되면 hidden_layer는 None x hiddenSize 크기를 가지는 행렬이고 W3은 hiddenSize x nbActions 크기를 가지는 행렬이므로 결과는 None x nbActions 크기를 가지는 행렬이 나오게 된다

이렇게 행렬곱을 하게되면 각각의 출력 레이어 노드에 들어가는 입력값들을 만들어낼 수 있다

이 값에 편향인 b3을 더해주면 각각의 출력 레이어 노드의 출력값들을 만들어낼 수 있다

```
In [38]: W3 = tf.Variable(tf.random.truncated_normal([hiddenSize, nbActions], stddev=1.0 / math.sqrt(float(hiddenSize)))) # 가중치
b3 = tf.Variable(tf.random.truncated_normal([nbActions], stddev=0.01)) # 편향
output_layer = tf.matmul(hidden_layer, W3) + b3 # 출력값
```

목표값 플레이스홀더

목표값 플레이스홀더 Y를 생성한다

Y는 데이터 유형이 float32이고 첫번째 차원의 수는 정해져있지 않고(가변적) 두번째 차원의 수는 nbActions의 값을 가지는 placeholder가 된다

```
In [39]: Y = tf.compat.v1.placeholder(tf.float32, [None, nbActions])
```

목표값과 출력값의 차이인 코스트

목표값과 출력값의 오차를 구하기 위해 여기서는 평균 제곱 오차를 사용한다

목표값(Y)에서 출력값(output_layer)을 뺀 다음 제곱을 해준다

그 다음 나온 값들을 모두 더해준 뒤 2 x batchSize로 나누어준다

batchSize는 한 번에 모델이 학습하는 데이터 샘플의 개수로 이렇게 2 x batchSize로 나누어주게 되면 코스트를 정규화해주게 된다

정규화를 통해 학습률과 batchSize에 따른 코스트 크기의 영향을 줄일 수 있다

```
In [40]: cost = tf.reduce_sum(tf.square(Y-output_layer)) / (2*batchSize)
```

경사하강법으로 코스트가 최소가 되는 값 찾기

경사하강법을 이용해 비용이 최소가 되는 값을 찾아준다

경사하강법을 사용할 때 학습률은 learningRate가 되고 minimize(cost)를 붙여주어 cost를 최소화하도록 설정해준다

```
In [41]: optimizer = tf.compat.v1.train.GradientDescentOptimizer(learningRate).minimize(cost)
```

랜덤값 함수

함수 randf는 s이상 e미만의 값을 가지는 float형태의 랜덤값을 리턴해준다

(float(random.randrange(0, (e - s) * 9999)) / 10000) + s; random.randrange는 0부터 (e - s) * 9999 미만의 랜덤값을 만들어준다

이 값에 float를 붙여 실수로 변환해준 다음 10000으로 나누어주면 0이상 (e - s)미만의 랜덤값이 나오게 된다

여기에 s를 더해주면 s이상 e미만의 랜덤값이 된다

```
In [42]: def randf(s, e):  
    return (float(random.randrange(0, (e - s) * 9999)) / 10000) + s;
```

환경 클래스

__init__ 함수

__init__ 함수는 초기화를 시켜준다

self.gridSize = gridSize

self.nbStates = self.gridSize * self.gridSize

self.state = np.empty(3, dtype = np.uint8)

np.empty는 초기화되지 않은 값으로 배열을 생성해주는 함수로 여기서는 길이가 3인 초기화되지 않은 배열이 생성된다

배열의 데이터 형식은 uint8(부호 없는 8비트 정수)가 된다

observe 함수

observe 함수는 화면정보를 리턴해준다 (관찰)

canvas = self.drawState()

drawState함수의 출력값을 canvas에 저장해준다

canvas = np.reshape(canvas, (-1,self.nbStates))

np.reshape는 배열을 재구성해주는 함수로 여기서는 canvas 배열을 self.nbStates개의 열을 가지는 배열로 재구성해준다(행의 개수는 -1로 설정되어 있으므로 자동으로 결정된다)
그 값을 다시 canvas에 저장해준다

return canvas

canvas를 리턴해준다

drawState 함수

drawState 함수는 블록과 바를 표시하여 화면정보를 리턴해준다

canvas = np.zeros((self.gridSize, self.gridSize))

모든 요소가 0으로 초기화된 self.gridSize x self.gridSize의 크기를 가지는 배열을 생성한다

canvas[self.state[0]-1, self.state[1]-1] = 1

self.state는 길이가 3인 배열로 0에는 블록의 행번호가 1에는 블록의 열번호를 저장하고 있다

self.state 값들을 계산할때는 gridSize를 기준으로 계산을 해 1 ~ gridSize이지만 실제로 나타낼 때는 0 ~ gridSize-1이기 때문에 각 값들에 -1을 해준 위치의 배열값을 1로 바꾸어준다

canvas[self.gridSize-1, self.state[2] -1 - 1] = 1

canvas[self.gridSize-1, self.state[2] -1] = 1

canvas[self.gridSize-1, self.state[2] -1 + 1] = 1

self.State[2]에는 바의 열번호를 저장하고 있다

따라서 바의 행번호는 self.gridSize-1(가장 마지막 행)을 해주고 열번호는 self.state[2]-1과 좌우 배열값들을 1로 바꾸어준다

return canvas

canvas를 리턴해준다

reset 함수

reset함수는 블록과 바의 위치를 리셋해준다

initialFruitColumn = random.randrange(1, self.gridSize + 1)

1이상 self.gridSize+1 미만의 랜덤값을 initialFruitColumn에 저장해준다

initialBucketPosition = random.randrange(2, self.gridSize + 1 - 1)

2이상 self.gridSize미만의 랜덤값을 initialBucketPosition에 저장해준다

바는 양옆에 픽셀까지 사용하기 때문에 블록과 다르게 2이상 self.gridSize미만으로 설정해준다

self.state = np.array([1, initialFruitColumn, initialBucketPosition])

블록의 시작은 항상 맨 위이어야 하기 때문에 self.state[0]은 1로 설정해주고 self.state[1]에는 블록의 열번호인 initialFruitColumn로 설정해주고 self.state[2]에는 바의 위치인 initialBucketPosition을 설정해준다(바는 항상 행번호가 마지막 행번호로 정해져있기 때문에 열번호만 변경해주면 된다)

return self.getState()

self.getState(현재 상태)를 리턴해준다

getState 함수

getState 함수는 현재 상태를 리턴해준다

stateInfo = self.state

stateInfo에 현재 state값을 불러온 뒤

fruit_row = stateInfo[0]

블럭의 행번호는 fruit_row에 저장

fruit_col = stateInfo[1]

블럭의 열번호는 fruit_col에 저장

basket = stateInfo[2]

바의 위치는 basket에 저장해준다

return fruit_row, fruit_col, basket

각각의 값들을 리턴해준다

getReward 함수

getReward 함수는 보상값을 리턴해준다

fruitRow, fruitColumn, basket = self.getState()

우선 getState 함수를 이용해 현재 상태를 가져온다

if (fruitRow == self.gridSize - 1): # 만약 블럭의 행번호가 self.gridSize - 1 과 같다면 (블럭이 가장 마지막 행까지 갔다면)

if (abs(fruitColumn - basket) <= 1): # 블럭의 열번호 - 바의 위치에 절대값이 <= 1이라면 (0이라면 == 위치가 같다면)

return 1 # 1을 리턴해준다

else:

return -1 # 위치가 같지 않다면 -1을 리턴해준다

else: # 블럭의 행번호가 gridSize - 1 과 같지 않다면 (블럭이 아직 가장 마지막 행까지 가지 않았다면)

return 0 # 0을 리턴해준다**

isGameOver 함수

isGameOver 함수는 게임이 끝났는지의 여부를 리턴해준다

if (self.state[0] == self.gridSize - 1): # 만약 블럭의 행번호가 self.girdSize-1과 같다면 (블럭이 가장 마지막 행까지 갔다면)

return True # True를 리턴해준다

else: # 블럭이 가장 마지막 행까지 가지 않았다면

return False # False를 리턴해준다

updateState 함수

updateState 함수는 action에 따라 바의 위치를 업데이트해주고 블럭의 위치를 업데이트해준다

if (action == 1): # action이 1이면 왼쪽으로 이동시켜준다

action = -1

elif (action == 2): # action이 2이면 그대로 있다

action = 0

else: # action이 1이나 2가 아니면 오른쪽으로 이동시켜준다

action = 1

fruitRow, fruitColumn, basket = self.getState()

getState함수를 이용해 현재상태를 가져온다

newBasket = min(max(2, basket + action), self.gridSize - 1) # 바 위치 변경

새로운 바의 위치는 우선 2와 basket + action 중 큰 것을 고른 다음 self.gridSize - 1과 비교하여 더 작은 값으로 설정해준다 (2이상 self.gridSize-1이하의 값으로 설정된다 바는 양옆의 픽셀까지 사용하기 때문에 2와 self.gridSize - 1로 설정해준다)

fruitRow = fruitRow + 1 # 블록을 아래로 이동

블록의 행번호도 하나 추가해준다

self.state = np.array([fruitRow, fruitColumn, newBasket])

state값을 새로운 값들로 업데이트 해준다

act 함수

act 함수는 행동을 수행한다

self.updateState(action)

updateState 함수를 사용해 바의 위치와 블록의 위치를 업데이트 시켜준다

reward = self.getReward()

getReward 함수를 사용해 보상값을 가져오고

gameOver = self.isGameOver()

isGameOver 함수를 사용해 게임이 끝났는지 확인한다

return self.observe(), reward, gameOver, self.getState()

앞에서 가져온 값들을 리턴해준다

```
In [43]: class CatchEnvironment():
# 초기화
def __init__(self, gridSize):
    self.gridSize = gridSize
    self.nbStates = self.gridSize * self.gridSize
    self.state = np.empty(3, dtype = np.uint8)

# 화면정보 리턴
def observe(self):
    canvas = self.drawState()
    canvas = np.reshape(canvas, (-1,self.nbStates))
    return canvas

# 블록과 바를 표시하여 화면정보 리턴
def drawState(self):
    canvas = np.zeros((self.gridSize, self.gridSize))
```

```

# 블럭 표시
canvas[self.state[0]-1, self.state[1]-1] = 1

# 바 표시
canvas[self.gridSize-1, self.state[2] -1 - 1] = 1
canvas[self.gridSize-1, self.state[2] -1] = 1
canvas[self.gridSize-1, self.state[2] -1 + 1] = 1
return canvas

# 블럭과 바 위치 초기화
def reset(self):
    initialFruitColumn = random.randrange(1, self.gridSize + 1)
    initialBucketPosition = random.randrange(2, self.gridSize + 1 - 1)
    self.state = np.array([1, initialFruitColumn, initialBucketPosition])
    return self.getState()

# 상태 리턴
def getState(self):
    stateInfo = self.state
    fruit_row = stateInfo[0]
    fruit_col = stateInfo[1]
    basket = stateInfo[2]
    return fruit_row, fruit_col, basket

# 보상값 리턴
def getReward(self):
    fruitRow, fruitColumn, basket = self.getState()
    if (fruitRow == self.gridSize - 1): # If the fruit has reached the bottom.
        if (abs(fruitColumn - basket) <= 1): # Check if the basket caught the fruit.
            return 1
        else:
            return -1
    else:
        return 0

# 게임오버 검사
def isGameOver(self):
    if (self.state[0] == self.gridSize - 1):
        return True
    else:
        return False

# 상태 업데이트
def updateState(self, action):
    if (action == 1):

```



```

        action = -1 # 왼쪽 이동
    elif (action == 2):
        action = 0 # 대기
    else:
        action = 1 # 오른쪽 이동
    fruitRow, fruitColumn, basket = self.getState()
    newBasket = min(max(2, basket + action), self.gridSize - 1) # 바 위치 변경
    fruitRow = fruitRow + 1 # 불럭을 아래로 이동
    self.state = np.array([fruitRow, fruitColumn, newBasket])

# 행동 수행 (1->왼쪽, 2->대기, 3->오른쪽)
def act(self, action):
    self.updateState(action)
    reward = self.getReward()
    gameOver = self.isGameOver()
    return self.observe(), reward, gameOver, self.getState()

```

메모리 클래스 (게임내용을 저장하고 나중에 배치로 묶어 학습에 사용)

__init__ 함수

__init__ 함수는 초기화를 시켜준다

```

self.maxMemory = maxMemory
self.gridSize = gridSize
self.nbStates = self.gridSize * self.gridSize
self.discount = discount
canvas = np.zeros((self.gridSize, self.gridSize))

```

canvas는 0으로 초기화된 self.gridSize x self.gridSize의 크기를 가지는 배열이 된다

```

canvas = np.reshape(canvas, (-1, self.nbStates))

```

위에서 생성한 canvas를 self.nbStates개의 열을 가지는 배열로 재배열해준다

```

self.inputState = np.empty((self.maxMemory, 100), dtype = np.float32)

```

self.inputState는 초기화가 되지 않은 self.maxMemory * 100의 크기를 가지는 배열이 된다

배열의 데이터 형식은 np.float32가 된다

```

self.actions = np.zeros(self.maxMemory, dtype = np.uint8)

```

self.actions는 0으로 초기화된 self.maxMemory의 길이를 가지는 배열이 된다 배열의 데이터 형식은 np.uint8이 된다

```

self.nextState = np.empty((self.maxMemory, 100), dtype = np.float32)

```

self.nextState는 초기화가 되지 않은 self.maxMemory x 100의 크기를 가지는 배열이 된다

배열의 데이터 형식은 np.float32가 된다

```

self.gameOver = np.empty(self.maxMemory, dtype = np.bool_)

```

self.gameOver는 초기화가 되지 않은 self.maxMemory의 길이를 가지는 배열이 된다

배열의 데이터 형식은 np.bool_이 된다

self.rewards = np.empty(self.maxMemory, dtype = np.int8)

self.rewards는 초기화가 되지 않은 self.maxMemory의 길이를 가지는 배열이 된다
배열의 데이터 형식은 np.int8이 된다

self.count = 0

self.current = 0

remember 함수

remember 함수는 경험을 ReplayMemory에 저장해준다

self.actions[self.current] = action

self.actions의 self.current번째 데이터는 action이 된다

self.rewards[self.current] = reward

self.rewards의 self.current번째 데이터는 reward가 된다

self.inputState[self.current, ...] = currentState

self.inputState의 self.current번째 행의 데이터는 currentState가 된다

self.nextState[self.current, ...] = nextState

self.nextState의 self.current번째 행의 데이터는 nextState가 된다

self.gameOver[self.current] = gameOver

self.gameOver의 self.current번째 데이터는 gameOver가 된다

self.count = max(self.count, self.current + 1)

self.count는 self.count와 self.current+1 중 더 큰 값이 된다

현재까지 사용된 데이터의 수를 업데이트해준다

self.current = (self.current + 1) % self.maxMemory

self.current는 self.current + 1을 self.maxMemory로 나눈 나머지값이 된다

이렇게 해주는 이유는 self.current + 1이 self.maxMemory 값과 같아질 경우 0으로 만들어 주기 때문이다
따라서 self.current는 0에서 self.maxMemory -1의 값을 가지게 된다

getBatch 함수

getBatch 함수는 저장된 게임내용에서 랜덤한 배치(batch)를 가져오는 함수입니다
배치(batch)는 한 번에 모델에 입력되는 데이터의 묶음을 말합니다

memoryLength = self.count

memoryLength는 self.count(현재까지 사용된 데이터의 수)가 된다

chosenBatchSize = min(batchSize, memoryLength)

chosenBatchSize는 batchSize와 memoryLength 중 더 작은 값이 된다

inputs = np.zeros((chosenBatchSize, nbStates))

inputs은 0으로 초기화된 chosenBatchSize x nbStates의 크기를 가지는 배열이 된다

```
targets = np.zeros((chosenBatchSize, nbActions))
```

targets은 0으로 초기화된 chosenBatchSize x nbActions의 크기를 가지는 배열이 된다

```
for i in range(chosenBatchSize): chosenBatchSize만큼 for문을 돌려준다
```

```
randomIndex = random.randrange(0, memoryLength)
```

randomIndex는 0이상 memoryLength 이하의 랜덤값을 가지게 된다

```
current_inputState = np.reshape(self.inputState[randomIndex], (1, 100))
```

current_inputState는 inputState의 randomIndex번째 행의 데이터를 1 x 100 크기의 배열로 재배열해준다

```
target = sess.run(model, feed_dict={X: current_inputState})
```

target은 tensorflow의 세션을 실행한 뒤 model을 설정해주고 X는 모델의 입력 플레이스홀더를 나타내고 current_inputState가 이 플레이스홀더에 입력이 된다

그 다음 모델을 실행하여 그 결과를 받을 수 있다

Q값을 불러온다

```
current_nextState = np.reshape(self.nextState[randomIndex], (1, 100))
```

current_nextState는 nextState의 randomIndex번째 행의 데이터를 1 x 100 크기의 배열로 재배열해준다

```
current_outputs = sess.run(model, feed_dict={X: current_nextState})
```

그 다음 target과 마찬가지로 세션을 실행한 다음 결과를 받는데 여기서는 플레이스홀더에 입력되는 값이 current_nextState로 바뀐다

```
nextStateMaxQ = np.amax(current_outputs)
```

nextStateMaxQ는 current_outputs에서 최대값을 가진다

다음 상태에서 가질 수 있는 가장 큰 Q값을 계산

```
if (self.gameOver[randomIndex] == True): # 만약 gameOver의 randomIndex번째 행의 값이 True라면 (게임이 끝났다면)
```

```
target[0, [self.actions[randomIndex]-1]] = self.rewards[randomIndex] # target의 첫번째 행의 self.actions의 randomIndex번째 데이터 - 1번째 열의 데이터는 self.rewards의 randomIndex번째 데이터가 된다
```

게임이 끝났을 때의 Q값은 reward(보상값)이 된다

```
else: # 만약 gameOver의 randomIndex 행의 값이 False라면 (게임이 끝나지 않았다면)
```

```
target[0, [self.actions[randomIndex]-1]] = self.rewards[randomIndex] +self.discount * nextStateMaxQ
```

True일때와 마찬가지로 self.reward의 데이터를 넣어주는데 이번에는 reward(보상값)에 self.discount(감소값) x nextStateMaxQ(가질 수 있는 가장 큰 Q값)을 더해준다

Q값을 업데이트 시켜주는 과정인데 Q-Learning의 업데이트 수식은 보상값 + 감소값 x maxQ(상태,행동)입니다

```
inputs[i] = current_inputState
```

```
targets[i] = target
```

```
return inputs, targets
```

for문이 끝나면 저장한 input과 targets을 리턴해준다

In [44]: **class** ReplayMemory:

```
# 초기화
```

```
def __init__(self, gridSize, maxMemory, discount):
```

```
    self.maxMemory = maxMemory
```

```
    self.gridSize = gridSize
```

```
    self.nbStates = self.gridSize * self.gridSize
```

```
    self.discount = discount
```

```
    canvas = np.zeros((self.gridSize, self.gridSize))
```

```

canvas = np.reshape(canvas, (-1,self.nbStates))
self.inputState = np.empty((self.maxMemory, 100), dtype = np.float32)
self.actions = np.zeros(self.maxMemory, dtype = np.uint8)
self.nextState = np.empty((self.maxMemory, 100), dtype = np.float32)
self.gameOver = np.empty(self.maxMemory, dtype = np.bool_)
self.rewards = np.empty(self.maxMemory, dtype = np.int8)
self.count = 0
self.current = 0

```

게임내용 추가

```

def remember(self, currentState, action, reward, nextState, gameOver):
    self.actions[self.current] = action
    self.rewards[self.current] = reward
    self.inputState[self.current, ...] = currentState
    self.nextState[self.current, ...] = nextState
    self.gameOver[self.current] = gameOver
    self.count = max(self.count, self.current + 1)
    self.current = (self.current + 1) % self.maxMemory

```

게임내용을 배치로 묶어서 리턴

```

def getBatch(self, model, batchSize, nbActions, nbStates, sess, X):
    memoryLength = self.count
    chosenBatchSize = min(batchSize, memoryLength)
    inputs = np.zeros((chosenBatchSize, nbStates))
    targets = np.zeros((chosenBatchSize, nbActions))

    for i in range(chosenBatchSize):
        # 메모리에서 랜덤하게 선택
        randomIndex = random.randrange(0, memoryLength)
        current_inputState = np.reshape(self.inputState[randomIndex], (1, 100))
        target = sess.run(model, feed_dict={X: current_inputState})

        current_nextState = np.reshape(self.nextState[randomIndex], (1, 100))
        current_outputs = sess.run(model, feed_dict={X: current_nextState})

        # 다음 상태의 최대 Q값
        nextStateMaxQ = np.amax(current_outputs)

        if (self.gameOver[randomIndex] == True):
            # 게임오버일때 Q값은 보상값으로 설정
            target[0, [self.actions[randomIndex]-1]] = self.rewards[randomIndex]
        else:
            # Q값을 계산
            # reward + discount(gamma) * max_a' Q(s',a')
            target[0, [self.actions[randomIndex]-1]] = self.rewards[randomIndex] +self.discount * nextStateMaxQ

```

```
        inputs[i] = current_inputState
        targets[i] = target
    return inputs, targets
```

메인함수

print("Training new model")

env = CatchEnvironment(gridSize)

env를 CatchEnvironment 클래스로 정의해준다

변수값으로 gridSize(게임화면 한 면의 픽셀 수)를 넘겨준다

memory = ReplayMemory(gridSize, maxMemory, discount)

memory를 ReplayMemory 클래스로 정의해준다

변수값으로 gridSize(게임화면 한 면의 픽셀 수), maxMemory(ReplayMemory 데이터의 최대크기), discount(감소값)를 넘겨준다

saver = tf.compat.v1.train.Saver()

모델과 파라미터를 저장하기 위해 saver를 설정해준다

winCount = 0

winCount는 0으로 초기화해준다

with tf.compat.v1.Session() as sess:

세션을 열고 사용한 후 자동으로 세션을 닫아주는 코드이다

여기서 sess는 세션을 나타내는 객체가 된다

sess.run(tf.compat.v1.global_variables_initializer()) 세션 내의 텐서플로우 변수들(tf.variables)을 초기화 시켜준다

for i in range(epoch):

epoch만큼 for문을 반복해준다

err = 0

한 번의 루프 동안 사용할 err 변수를 0으로 초기화 해준다

env.reset()

reset 함수를 사용하여 블록과 바의 위치를 리셋해준다

isGameOver = False

isGameOver 변수의 루프 초기화 값을 False로 설정해준다

currentState = env.observe()

currentState는 observe 함수를 사용하여 현재 화면의 상태를 가져온다

while (isGameOver != True):

만약 게임이 끝나지 않았다면 (만약 게임이 끝난다면 while문이 풀리고 다시 for문으로 돌아가 게임을 재시작하게 된다)

action = -9999

action을 -9999로 설정해주고(초기화)

global epsilon

전역변수인 epsilon을 불러온다

epsilon은 랜덤하게 행동할 확률을 말한다

epsilon greedy 기법에 따라 action을 취하게 된다

if (randf(0, 1) <= epsilon):

randf 함수를 이용하여 float 형태의 0이상 1미만인 랜덤값을 가져온 뒤 epsilon과 비교하여 epsilon이 같거나 더 크다면(epsilon의 확률만큼 랜덤한 행동을 합니다)

action = random.randrange(1, nbActions+1)

action은 1이상 nbActions + 1 미만인 랜덤값을 가지게 된다

else:

반대로 epsilon 보다 랜덤값이 작게 나온다면 (1-epsilon의 확률만큼 greedy 기법에 따라 Q값을 기반으로 가장 높은 가치를 가진 행동을 선택한다)

q = sess.run(output_layer, feed_dict={X: currentState})

위에서 정의했던 딥러닝에 currentState를 입력값으로 넣어 돌려 q에 넣어준다

index = q.argmax()

action = index + 1

딥러닝을 돌려 얻은 q값들 중 가장 큰 값에 1을 더하여 action으로 설정해준다

if (epsilon > epsilonMinimumValue):

만약 epsilon이 epsilonMinimumValue(epsilon의 최소값)보다 크다면

epsilon = epsilon * 0.999

epsilon의 값을 감소시켜준다 (랜덤으로 행동할 확률을 감소시킨다)

이렇게 되면 학습 초반에는 랜덤으로 행동할 확률이 높고 학습이 진행될 수록 랜덤으로 행동할 확률이 줄어든다

nextState, reward, gameOver, stateInfo = env.act(action)

act 함수를 사용해 nextState, reward, gameOver, stateInfo를 얻어낸다

act 함수를 사용하게 되면 우선 updateState 함수를 사용해 상태를 업데이트 해준다

이때 updateState에 아까 얻은 action값을 넣어주게 된다

nextState: observe 함수를 사용해 얻은 값으로 화면 정보를 리턴받는다

reward: getReward 함수를 사용해 보상값을 리턴받는다

gameOver: isGameOver 함수를 사용하여 게임의 종료 여부를 확인한다

stateInfo: getState 함수를 사용하여 게임의 상태를 리턴받는다

if (reward == 1): 만약 reward가 1이라면 (블럭이 마지막 행까지 갔을 때 바와 위치가 같다면(게임을 이겼다면))

winCount = winCount + 1

winCount를 1 올려준다

memory.remember(currentState, action, reward, nextState, gameOver)

remember 함수를 사용해 경험을 Replay Memory에 저장해준다

currentState = nextState

현재 상태를 다음 상태(act함수를 실행한)로 업데이트 해준다

isGameOver = gameOver

게임 종료 여부를 업데이트 시켜준다

inputs, targets = memory.getBatch(output_layer, batchSize, nbActions, nbStates, sess, X)

getBatch 함수를 사용하여 Replay Memory에서 학습에 사용할 batch 데이터를 불러온다

_, loss = sess.run([optimizer, cost], feed_dict={X: inputs, Y: targets})

세션에서 optimizer(경사하강법을 이용한 최적값)과 cost(오차)를 실행해 optimizer을 이용해 최적화된 cost를 만들어낸다

가져온 batch 데이터를 이용해 입력값은 inputs을 타겟값은 targets로 설정한다

optimizer 결과는 사용하지 않기 때문에 _로 표시하고 cost 결과는 loss에 저장해준다

err = err + loss

err에 위에서 나온 결과인 loss를 더해준다

****print("Epoch " + str(i) + ": err = " + str(err) + ": Win count = " + str(winCount) + " Win ratio = " + str(float(winCount)/float(i+1)*100))****

for문이 돌때마다 각각의 값들을 출력해준다

Win ratio는 현재 winCount를 i+1(i는 0부터 시작하기 때문에 1을 더해준다)를 나눈값에 100을 곱해 나타내준다

save_path = saver.save(sess, 'model')

만들어진 모델을 model이라는 이름으로 저장해준다 (현재 파일이 있는 경로에 저장된다)

print("Finish")

```
In [45]: def main(_):
    print("Training new model")

    # 환경 정의
    env = CatchEnvironment(gridSize)

    # 메모리 정의
    memory = ReplayMemory(gridSize, maxMemory, discount)

    # 세이버 설정
    saver = tf.compat.v1.train.Saver()

    winCount = 0
    with tf.compat.v1.Session() as sess:
        sess.run(tf.compat.v1.global_variables_initializer())

        for i in range(epoch):
            err = 0
            env.reset()

            isGameOver = False

            currentState = env.observe()

            while (isGameOver != True):
                action = -9999

                # 랜덤으로 행동을 할지 Q값에 따라 행동할지 결정
```

```

global epsilon
if (randf(0, 1) <= epsilon):
    action = random.randrange(1, nbActions+1)
else:
    q = sess.run(output_layer, feed_dict={X: currentState})
    index = q.argmax()
    action = index + 1

# 랜덤으로 행동할 확률 감소
if (epsilon > epsilonMinimumValue):
    epsilon = epsilon * 0.999

# 행동 수행
nextState, reward, gameOver, stateInfo = env.act(action)

# 승리 횟수 설정
if (reward == 1):
    winCount = winCount + 1

# 메모리에 저장
memory.remember(currentState, action, reward, nextState, gameOver)

# 다음 상태 설정
currentState = nextState
isGameOver = gameOver

# 입력과 출력 데이터 배치를 구함
inputs, targets = memory.getBatch(output_layer, batchSize, nbActions, nbStates, sess, X)

# 학습 수행
_, loss = sess.run([optimizer, cost], feed_dict={X: inputs, Y: targets})
err = err + loss

print("Epoch " + str(i) + ": err = " + str(err) + ": Win count = " + str(winCount) + " Win ratio = " + str(float(winCount)/float(i+1)*100))

# 모델 세션 저장
save_path = saver.save(sess, 'model')

print("Finish")

```

메인 함수 실행

```
In [46]: if __name__ == '__main__':
```


tf.compat.v1.app.run()

Training new model

Epoch 0: err = 0.006210632644069847: Win count = 0 Win ratio = 0.0
Epoch 1: err = 0.13175524293910712: Win count = 0 Win ratio = 0.0
Epoch 2: err = 0.17373778542969376: Win count = 0 Win ratio = 0.0
Epoch 3: err = 0.20777934463694692: Win count = 0 Win ratio = 0.0
Epoch 4: err = 0.20776284579187632: Win count = 0 Win ratio = 0.0
Epoch 5: err = 0.32736863382160664: Win count = 0 Win ratio = 0.0
Epoch 6: err = 0.35962055902928114: Win count = 0 Win ratio = 0.0
Epoch 7: err = 0.3084113756194711: Win count = 0 Win ratio = 0.0
Epoch 8: err = 0.2549772076308727: Win count = 0 Win ratio = 0.0
Epoch 9: err = 0.3277985155582428: Win count = 0 Win ratio = 0.0
Epoch 10: err = 0.3028283454477787: Win count = 0 Win ratio = 0.0
Epoch 11: err = 0.2646270878612995: Win count = 0 Win ratio = 0.0
Epoch 12: err = 0.2169065224006772: Win count = 0 Win ratio = 0.0
Epoch 13: err = 0.2877187002450228: Win count = 0 Win ratio = 0.0
Epoch 14: err = 0.22772240359336138: Win count = 0 Win ratio = 0.0
Epoch 15: err = 0.23943635076284409: Win count = 0 Win ratio = 0.0
Epoch 16: err = 0.23223904985934496: Win count = 0 Win ratio = 0.0
Epoch 17: err = 0.20301990769803524: Win count = 0 Win ratio = 0.0
Epoch 18: err = 0.17319162003695965: Win count = 1 Win ratio = 5.263157894736842
Epoch 19: err = 0.2352193035185337: Win count = 1 Win ratio = 5.0
Epoch 20: err = 0.1823448585346341: Win count = 2 Win ratio = 9.523809523809524
Epoch 21: err = 0.18840092606842518: Win count = 2 Win ratio = 9.090909090909092
Epoch 22: err = 0.2667874824255705: Win count = 2 Win ratio = 8.695652173913043
Epoch 23: err = 0.17130382731556892: Win count = 2 Win ratio = 8.333333333333332
Epoch 24: err = 0.23114200960844755: Win count = 2 Win ratio = 8.0
Epoch 25: err = 0.20083474554121494: Win count = 3 Win ratio = 11.538461538461538
Epoch 26: err = 0.1765302335843444: Win count = 3 Win ratio = 11.111111111111111
Epoch 27: err = 0.25021445006132126: Win count = 3 Win ratio = 10.714285714285714
Epoch 28: err = 0.25319355353713036: Win count = 3 Win ratio = 10.344827586206897
Epoch 29: err = 0.24055784195661545: Win count = 4 Win ratio = 13.333333333333334
Epoch 30: err = 0.27754813712090254: Win count = 4 Win ratio = 12.903225806451612
Epoch 31: err = 0.36915131472051144: Win count = 5 Win ratio = 15.625
Epoch 32: err = 0.2559082880616188: Win count = 6 Win ratio = 18.181818181818183
Epoch 33: err = 0.2582416255027056: Win count = 7 Win ratio = 20.588235294117645
Epoch 34: err = 0.30837620981037617: Win count = 8 Win ratio = 22.857142857142858
Epoch 35: err = 0.328916946426034: Win count = 8 Win ratio = 22.222222222222222
Epoch 36: err = 0.3105590995401144: Win count = 8 Win ratio = 21.62162162162162
Epoch 37: err = 0.3593046348541975: Win count = 8 Win ratio = 21.052631578947366
Epoch 38: err = 0.38076117634773254: Win count = 8 Win ratio = 20.51282051282051
Epoch 39: err = 0.3931809738278389: Win count = 8 Win ratio = 20.0
Epoch 40: err = 0.3067593090236187: Win count = 8 Win ratio = 19.51219512195122
Epoch 41: err = 0.2710862476378679: Win count = 9 Win ratio = 21.428571428571427
Epoch 42: err = 0.3426066655665636: Win count = 9 Win ratio = 20.930232558139537
Epoch 43: err = 0.2849228587001562: Win count = 9 Win ratio = 20.454545454545457
Epoch 44: err = 0.2751770056784153: Win count = 10 Win ratio = 22.222222222222222
Epoch 45: err = 0.303482370916754: Win count = 11 Win ratio = 23.91304347826087
Epoch 46: err = 0.34951798617839813: Win count = 11 Win ratio = 23.404255319148938
Epoch 47: err = 0.4282461702823639: Win count = 11 Win ratio = 22.916666666666664
Epoch 48: err = 0.34239001385867596: Win count = 11 Win ratio = 22.448979591836736
Epoch 49: err = 0.33551312796771526: Win count = 12 Win ratio = 24.0
Epoch 50: err = 0.3068511765450239: Win count = 13 Win ratio = 25.49019607843137
Epoch 51: err = 0.38515353202819824: Win count = 13 Win ratio = 25.0
Epoch 52: err = 0.24145356845110655: Win count = 13 Win ratio = 24.528301886792452
Epoch 53: err = 0.2866846350952983: Win count = 13 Win ratio = 24.074074074074073
Epoch 54: err = 0.3425815459340811: Win count = 13 Win ratio = 23.636363636363637
Epoch 55: err = 0.3339632200077176: Win count = 13 Win ratio = 23.214285714285715
Epoch 56: err = 0.32057400699704885: Win count = 14 Win ratio = 24.561403508771928
Epoch 57: err = 0.25963449850678444: Win count = 14 Win ratio = 24.137931034482758
Epoch 58: err = 0.29080385714769363: Win count = 15 Win ratio = 25.423728813559322

Epoch 59: err = 0.37676159478724003: Win count = 16 Win ratio = 26.666666666666668
Epoch 60: err = 0.32168901711702347: Win count = 17 Win ratio = 27.86885245901639
Epoch 61: err = 0.32402967289090157: Win count = 17 Win ratio = 27.419354838709676
Epoch 62: err = 0.3118943180888891: Win count = 17 Win ratio = 26.984126984126984
Epoch 63: err = 0.3278345298022032: Win count = 17 Win ratio = 26.5625
Epoch 64: err = 0.2643015021458268: Win count = 17 Win ratio = 26.153846153846157
Epoch 65: err = 0.32341026328504086: Win count = 18 Win ratio = 27.27272727272727
Epoch 66: err = 0.3291377127170563: Win count = 19 Win ratio = 28.35820895522388
Epoch 67: err = 0.41750209871679544: Win count = 20 Win ratio = 29.411764705882355
Epoch 68: err = 0.30872078984975815: Win count = 20 Win ratio = 28.985507246376812
Epoch 69: err = 0.3336247345432639: Win count = 21 Win ratio = 30.0
Epoch 70: err = 0.42509228736162186: Win count = 21 Win ratio = 29.577464788732392
Epoch 71: err = 0.3848083224147558: Win count = 22 Win ratio = 30.555555555555557
Epoch 72: err = 0.26648691669106483: Win count = 22 Win ratio = 30.136986301369863
Epoch 73: err = 0.401422331109643: Win count = 22 Win ratio = 29.72972972972973
Epoch 74: err = 0.3119311220943928: Win count = 23 Win ratio = 30.666666666666664
Epoch 75: err = 0.4071196559816599: Win count = 23 Win ratio = 30.263157894736842
Epoch 76: err = 0.41617878153920174: Win count = 23 Win ratio = 29.87012987012987
Epoch 77: err = 0.4369244873523712: Win count = 24 Win ratio = 30.76923076923077
Epoch 78: err = 0.2900168988853693: Win count = 25 Win ratio = 31.645569620253166
Epoch 79: err = 0.42095207422971725: Win count = 25 Win ratio = 31.25
Epoch 80: err = 0.2906584534794092: Win count = 25 Win ratio = 30.864197530864196
Epoch 81: err = 0.3119247332215309: Win count = 26 Win ratio = 31.70731707317073
Epoch 82: err = 0.45257358625531197: Win count = 26 Win ratio = 31.32530120481928
Epoch 83: err = 0.33499860018491745: Win count = 26 Win ratio = 30.952380952380953
Epoch 84: err = 0.2183801420032978: Win count = 26 Win ratio = 30.58823529411765
Epoch 85: err = 0.27814305759966373: Win count = 27 Win ratio = 31.3953488372093
Epoch 86: err = 0.2561373598873615: Win count = 27 Win ratio = 31.03448275862069
Epoch 87: err = 0.3788217566907406: Win count = 28 Win ratio = 31.818181818181817
Epoch 88: err = 0.30276817828416824: Win count = 28 Win ratio = 31.46067415730337
Epoch 89: err = 0.3114276872947812: Win count = 29 Win ratio = 32.22222222222222
Epoch 90: err = 0.3556164335459471: Win count = 29 Win ratio = 31.868131868131865
Epoch 91: err = 0.37806474417448044: Win count = 30 Win ratio = 32.608695652173914
Epoch 92: err = 0.35263350047171116: Win count = 30 Win ratio = 32.25806451612903
Epoch 93: err = 0.32070532999932766: Win count = 31 Win ratio = 32.97872340425532
Epoch 94: err = 0.2789298091083765: Win count = 31 Win ratio = 32.631578947368425
Epoch 95: err = 0.32409244775772095: Win count = 31 Win ratio = 32.291666666666667
Epoch 96: err = 0.4037699420005083: Win count = 32 Win ratio = 32.98969072164948
Epoch 97: err = 0.2858038078993559: Win count = 32 Win ratio = 32.6530612244898
Epoch 98: err = 0.26248266734182835: Win count = 33 Win ratio = 33.33333333333333
Epoch 99: err = 0.21032274886965752: Win count = 33 Win ratio = 33.0
Epoch 100: err = 0.3074294738471508: Win count = 34 Win ratio = 33.663366336633665
Epoch 101: err = 0.3008042313158512: Win count = 34 Win ratio = 33.33333333333333
Epoch 102: err = 0.27891735546290874: Win count = 34 Win ratio = 33.00970873786408
Epoch 103: err = 0.2923729922622442: Win count = 34 Win ratio = 32.69230769230769
Epoch 104: err = 0.26842666417360306: Win count = 35 Win ratio = 33.33333333333333
Epoch 105: err = 0.24735314771533012: Win count = 35 Win ratio = 33.0188679245283
Epoch 106: err = 0.30662233009934425: Win count = 35 Win ratio = 32.71028037383177
Epoch 107: err = 0.28870732337236404: Win count = 35 Win ratio = 32.407407407407405
Epoch 108: err = 0.33229823410511017: Win count = 36 Win ratio = 33.02752293577982
Epoch 109: err = 0.28115597553551197: Win count = 36 Win ratio = 32.72727272727273
Epoch 110: err = 0.36406139843165874: Win count = 36 Win ratio = 32.432432432432435
Epoch 111: err = 0.32611602544784546: Win count = 37 Win ratio = 33.035714285714285
Epoch 112: err = 0.3743466604501009: Win count = 37 Win ratio = 32.743362831858406
Epoch 113: err = 0.28737848810851574: Win count = 37 Win ratio = 32.45614035087719
Epoch 114: err = 0.29839761927723885: Win count = 38 Win ratio = 33.04347826086956
Epoch 115: err = 0.30406877864152193: Win count = 39 Win ratio = 33.62068965517241
Epoch 116: err = 0.23837811313569546: Win count = 40 Win ratio = 34.18803418803419
Epoch 117: err = 0.25108332745730877: Win count = 40 Win ratio = 33.89830508474576
Epoch 118: err = 0.26326456293463707: Win count = 41 Win ratio = 34.45378151260504

Epoch 119: err = 0.23013906832784414: Win count = 42 Win ratio = 35.0
Epoch 120: err = 0.3194756470620632: Win count = 42 Win ratio = 34.710743801652896
Epoch 121: err = 0.18980211578309536: Win count = 42 Win ratio = 34.42622950819672
Epoch 122: err = 0.30144900269806385: Win count = 42 Win ratio = 34.146341463414636
Epoch 123: err = 0.18371485825628042: Win count = 42 Win ratio = 33.87096774193548
Epoch 124: err = 0.26397733204066753: Win count = 43 Win ratio = 34.4
Epoch 125: err = 0.22464224137365818: Win count = 43 Win ratio = 34.12698412698413
Epoch 126: err = 0.2952934578061104: Win count = 43 Win ratio = 33.85826771653544
Epoch 127: err = 0.26608058623969555: Win count = 43 Win ratio = 33.59375
Epoch 128: err = 0.17584886122494936: Win count = 43 Win ratio = 33.33333333333333
Epoch 129: err = 0.2498998511582613: Win count = 43 Win ratio = 33.07692307692307
Epoch 130: err = 0.26905602402985096: Win count = 43 Win ratio = 32.82442748091603
Epoch 131: err = 0.21448140684515238: Win count = 43 Win ratio = 32.57575757575758
Epoch 132: err = 0.22890938445925713: Win count = 43 Win ratio = 32.33082706766917
Epoch 133: err = 0.2715856544673443: Win count = 44 Win ratio = 32.83582089552239
Epoch 134: err = 0.24879213981330395: Win count = 44 Win ratio = 32.592592592592595
Epoch 135: err = 0.21946069691330194: Win count = 45 Win ratio = 33.088235294117645
Epoch 136: err = 0.22678588144481182: Win count = 45 Win ratio = 32.846715328467155
Epoch 137: err = 0.2083324734121561: Win count = 45 Win ratio = 32.608695652173914
Epoch 138: err = 0.24677791818976402: Win count = 46 Win ratio = 33.093525179856115
Epoch 139: err = 0.17263842653483152: Win count = 47 Win ratio = 33.57142857142857
Epoch 140: err = 0.22582961525768042: Win count = 47 Win ratio = 33.33333333333333
Epoch 141: err = 0.20297237113118172: Win count = 48 Win ratio = 33.80281690140845
Epoch 142: err = 0.18308023922145367: Win count = 48 Win ratio = 33.56643356643357
Epoch 143: err = 0.18765699490904808: Win count = 48 Win ratio = 33.33333333333333
Epoch 144: err = 0.1865274552255869: Win count = 48 Win ratio = 33.10344827586207
Epoch 145: err = 0.19601295609027147: Win count = 48 Win ratio = 32.87671232876712
Epoch 146: err = 0.21598275750875473: Win count = 49 Win ratio = 33.33333333333333
Epoch 147: err = 0.20041605550795794: Win count = 49 Win ratio = 33.108108108108105
Epoch 148: err = 0.16525801550596952: Win count = 49 Win ratio = 32.88590604026846
Epoch 149: err = 0.18392590153962374: Win count = 49 Win ratio = 32.666666666666664
Epoch 150: err = 0.17120049614459276: Win count = 49 Win ratio = 32.450331125827816
Epoch 151: err = 0.1741490215063095: Win count = 50 Win ratio = 32.89473684210527
Epoch 152: err = 0.16638302709907293: Win count = 50 Win ratio = 32.6797385620915
Epoch 153: err = 0.1436861352995038: Win count = 51 Win ratio = 33.116883116883116
Epoch 154: err = 0.17530770786106586: Win count = 51 Win ratio = 32.903225806451616
Epoch 155: err = 0.16315885353833437: Win count = 51 Win ratio = 32.69230769230769
Epoch 156: err = 0.15971041098237038: Win count = 52 Win ratio = 33.12101910828025
Epoch 157: err = 0.13449757546186447: Win count = 52 Win ratio = 32.91139240506329
Epoch 158: err = 0.1754562295973301: Win count = 53 Win ratio = 33.33333333333333
Epoch 159: err = 0.16237993072718382: Win count = 53 Win ratio = 33.125
Epoch 160: err = 0.1702138027176261: Win count = 54 Win ratio = 33.54037267080746
Epoch 161: err = 0.1675199307501316: Win count = 55 Win ratio = 33.95061728395062
Epoch 162: err = 0.15645254123955965: Win count = 55 Win ratio = 33.74233128834356
Epoch 163: err = 0.14056820888072252: Win count = 55 Win ratio = 33.53658536585366
Epoch 164: err = 0.15535920951515436: Win count = 56 Win ratio = 33.939393939393945
Epoch 165: err = 0.1248756144195795: Win count = 56 Win ratio = 33.734939759036145
Epoch 166: err = 0.13907057512551546: Win count = 56 Win ratio = 33.532934131736525
Epoch 167: err = 0.1546574318781495: Win count = 56 Win ratio = 33.33333333333333
Epoch 168: err = 0.14209611155092716: Win count = 57 Win ratio = 33.72781065088758
Epoch 169: err = 0.10584326460957527: Win count = 57 Win ratio = 33.52941176470588
Epoch 170: err = 0.11776558216661215: Win count = 57 Win ratio = 33.33333333333333
Epoch 171: err = 0.14643446821719408: Win count = 58 Win ratio = 33.72093023255814
Epoch 172: err = 0.12475594971328974: Win count = 59 Win ratio = 34.104046242774565
Epoch 173: err = 0.15432651899755: Win count = 59 Win ratio = 33.90804597701149
Epoch 174: err = 0.10654324293136597: Win count = 60 Win ratio = 34.285714285714285
Epoch 175: err = 0.10728025063872337: Win count = 60 Win ratio = 34.090909090909086
Epoch 176: err = 0.12934613693505526: Win count = 60 Win ratio = 33.89830508474576
Epoch 177: err = 0.1508363112807274: Win count = 61 Win ratio = 34.26966292134831
Epoch 178: err = 0.14541932474821806: Win count = 62 Win ratio = 34.63687150837989

Epoch 179: err = 0.14788292907178402: Win count = 62 Win ratio = 34.44444444444444
Epoch 180: err = 0.11198484245687723: Win count = 62 Win ratio = 34.25414364640884
Epoch 181: err = 0.1720833834260702: Win count = 62 Win ratio = 34.065934065934066
Epoch 182: err = 0.13024374563246965: Win count = 62 Win ratio = 33.87978142076503
Epoch 183: err = 0.09808824816718698: Win count = 63 Win ratio = 34.23913043478261
Epoch 184: err = 0.09297986561432481: Win count = 63 Win ratio = 34.054054054054056
Epoch 185: err = 0.15895139798521996: Win count = 63 Win ratio = 33.87096774193548
Epoch 186: err = 0.10321745369583368: Win count = 64 Win ratio = 34.22459893048128
Epoch 187: err = 0.1093177730217576: Win count = 64 Win ratio = 34.04255319148936
Epoch 188: err = 0.1420575324445963: Win count = 64 Win ratio = 33.86243386243386
Epoch 189: err = 0.12392456829547882: Win count = 65 Win ratio = 34.21052631578947
Epoch 190: err = 0.10449209809303284: Win count = 65 Win ratio = 34.031413612565444
Epoch 191: err = 0.09858974581584334: Win count = 65 Win ratio = 33.85416666666667
Epoch 192: err = 0.08631174173206091: Win count = 66 Win ratio = 34.196891191709845
Epoch 193: err = 0.10048008896410465: Win count = 66 Win ratio = 34.02061855670103
Epoch 194: err = 0.11601757165044546: Win count = 66 Win ratio = 33.84615384615385
Epoch 195: err = 0.08075416879728436: Win count = 66 Win ratio = 33.6734693877551
Epoch 196: err = 0.11340955970808864: Win count = 66 Win ratio = 33.50253807106599
Epoch 197: err = 0.09728623647242785: Win count = 66 Win ratio = 33.33333333333333
Epoch 198: err = 0.08605586597695947: Win count = 67 Win ratio = 33.66834170854271
Epoch 199: err = 0.08290230762213469: Win count = 67 Win ratio = 33.5
Epoch 200: err = 0.11254832521080971: Win count = 67 Win ratio = 33.33333333333333
Epoch 201: err = 0.11941061541438103: Win count = 67 Win ratio = 33.16831683168317
Epoch 202: err = 0.08841582341119647: Win count = 67 Win ratio = 33.004926108374384
Epoch 203: err = 0.10361084528267384: Win count = 68 Win ratio = 33.33333333333333
Epoch 204: err = 0.08846197696402669: Win count = 68 Win ratio = 33.170731707317074
Epoch 205: err = 0.10963775031268597: Win count = 68 Win ratio = 33.00970873786408
Epoch 206: err = 0.10722731286659837: Win count = 69 Win ratio = 33.33333333333333
Epoch 207: err = 0.11826739460229874: Win count = 70 Win ratio = 33.65384615384615
Epoch 208: err = 0.09756380598992109: Win count = 71 Win ratio = 33.97129186602871
Epoch 209: err = 0.09970212308689952: Win count = 72 Win ratio = 34.285714285714285
Epoch 210: err = 0.0896970983594656: Win count = 72 Win ratio = 34.12322274881517
Epoch 211: err = 0.11373890098184347: Win count = 72 Win ratio = 33.9622641509434
Epoch 212: err = 0.12616251967847347: Win count = 73 Win ratio = 34.27230046948357
Epoch 213: err = 0.11496576853096485: Win count = 74 Win ratio = 34.57943925233645
Epoch 214: err = 0.10464883083477616: Win count = 74 Win ratio = 34.418604651162795
Epoch 215: err = 0.10180452326312661: Win count = 74 Win ratio = 34.25925925925926
Epoch 216: err = 0.08428150555118918: Win count = 74 Win ratio = 34.10138248847927
Epoch 217: err = 0.12886298168450594: Win count = 74 Win ratio = 33.94495412844037
Epoch 218: err = 0.11287527997046709: Win count = 75 Win ratio = 34.24657534246575
Epoch 219: err = 0.09941821312531829: Win count = 75 Win ratio = 34.090909090909086
Epoch 220: err = 0.07556593604385853: Win count = 75 Win ratio = 33.93665158371041
Epoch 221: err = 0.11403690790757537: Win count = 75 Win ratio = 33.78378378378378
Epoch 222: err = 0.08713729586452246: Win count = 75 Win ratio = 33.6322869955157
Epoch 223: err = 0.09655537642538548: Win count = 76 Win ratio = 33.92857142857143
Epoch 224: err = 0.08951384527608752: Win count = 76 Win ratio = 33.77777777777778
Epoch 225: err = 0.10579898813739419: Win count = 76 Win ratio = 33.6283185840708
Epoch 226: err = 0.0853451956063509: Win count = 76 Win ratio = 33.480176211453745
Epoch 227: err = 0.07428906159475446: Win count = 76 Win ratio = 33.33333333333333
Epoch 228: err = 0.09252736391499639: Win count = 77 Win ratio = 33.624454148471614
Epoch 229: err = 0.09525833325460553: Win count = 78 Win ratio = 33.91304347826087
Epoch 230: err = 0.11515199532732368: Win count = 79 Win ratio = 34.1991341991342
Epoch 231: err = 0.10660690208896995: Win count = 79 Win ratio = 34.05172413793103
Epoch 232: err = 0.11772457929328084: Win count = 79 Win ratio = 33.90557939914164
Epoch 233: err = 0.08534001000225544: Win count = 79 Win ratio = 33.76068376068376
Epoch 234: err = 0.12550850864499807: Win count = 79 Win ratio = 33.61702127659574
Epoch 235: err = 0.10498769208788872: Win count = 79 Win ratio = 33.47457627118644
Epoch 236: err = 0.09221250703558326: Win count = 79 Win ratio = 33.33333333333333
Epoch 237: err = 0.09345626831054688: Win count = 80 Win ratio = 33.61344537815126
Epoch 238: err = 0.08279039897024632: Win count = 81 Win ratio = 33.89121338912134

Epoch 239: err = 0.07783429091796279: Win count = 82 Win ratio = 34.166666666666664
Epoch 240: err = 0.07206739904358983: Win count = 83 Win ratio = 34.439834024896264
Epoch 241: err = 0.09305571904405951: Win count = 84 Win ratio = 34.710743801652896
Epoch 242: err = 0.07219009846448898: Win count = 84 Win ratio = 34.5679012345679
Epoch 243: err = 0.09049400221556425: Win count = 85 Win ratio = 34.83606557377049
Epoch 244: err = 0.08508707163855433: Win count = 85 Win ratio = 34.69387755102041
Epoch 245: err = 0.08200961537659168: Win count = 85 Win ratio = 34.552845528455286
Epoch 246: err = 0.09690146148204803: Win count = 86 Win ratio = 34.81781376518219
Epoch 247: err = 0.06844395981170237: Win count = 86 Win ratio = 34.67741935483871
Epoch 248: err = 0.1074860654771328: Win count = 87 Win ratio = 34.93975903614458
Epoch 249: err = 0.07921492541208863: Win count = 87 Win ratio = 34.8
Epoch 250: err = 0.0862318598665297: Win count = 88 Win ratio = 35.0597609561753
Epoch 251: err = 0.07493132213130593: Win count = 88 Win ratio = 34.92063492063492
Epoch 252: err = 0.09036194393411279: Win count = 89 Win ratio = 35.177865612648226
Epoch 253: err = 0.07687456347048283: Win count = 89 Win ratio = 35.039370078740156
Epoch 254: err = 0.07706682523712516: Win count = 89 Win ratio = 34.90196078431372
Epoch 255: err = 0.07295549381524324: Win count = 89 Win ratio = 34.765625
Epoch 256: err = 0.09638597443699837: Win count = 89 Win ratio = 34.63035019455253
Epoch 257: err = 0.09646324953064322: Win count = 89 Win ratio = 34.49612403100775
Epoch 258: err = 0.049546714406460524: Win count = 89 Win ratio = 34.36293436293436
Epoch 259: err = 0.13010309264063835: Win count = 89 Win ratio = 34.23076923076923
Epoch 260: err = 0.1520265443250537: Win count = 89 Win ratio = 34.099616858237546
Epoch 261: err = 0.11097892466932535: Win count = 90 Win ratio = 34.35114503816794
Epoch 262: err = 0.09930208791047335: Win count = 91 Win ratio = 34.60076045627377
Epoch 263: err = 0.10238214815035462: Win count = 92 Win ratio = 34.84848484848485
Epoch 264: err = 0.09524346049875021: Win count = 93 Win ratio = 35.094339622641506
Epoch 265: err = 0.20032376144081354: Win count = 94 Win ratio = 35.338345864661655
Epoch 266: err = 0.07492705201730132: Win count = 94 Win ratio = 35.2059925093633
Epoch 267: err = 0.08799226814880967: Win count = 95 Win ratio = 35.44776119402985
Epoch 268: err = 0.0724543365649879: Win count = 96 Win ratio = 35.687732342007436
Epoch 269: err = 0.07825406640768051: Win count = 97 Win ratio = 35.92592592592593
Epoch 270: err = 0.10586467757821083: Win count = 97 Win ratio = 35.79335793357934
Epoch 271: err = 0.0982293775305152: Win count = 98 Win ratio = 36.029411764705884
Epoch 272: err = 0.16584003530442715: Win count = 99 Win ratio = 36.26373626373626
Epoch 273: err = 0.19185317400842905: Win count = 100 Win ratio = 36.496350364963504
Epoch 274: err = 0.11368102114647627: Win count = 100 Win ratio = 36.36363636363637
Epoch 275: err = 0.12338806828483939: Win count = 101 Win ratio = 36.594202898550726
Epoch 276: err = 0.06881731515750289: Win count = 102 Win ratio = 36.8231046931408
Epoch 277: err = 0.12142746243625879: Win count = 102 Win ratio = 36.69064748201439
Epoch 278: err = 0.12638039514422417: Win count = 102 Win ratio = 36.55913978494624
Epoch 279: err = 0.09443385526537895: Win count = 103 Win ratio = 36.78571428571429
Epoch 280: err = 0.08873860538005829: Win count = 104 Win ratio = 37.01067615658363
Epoch 281: err = 0.13481033500283957: Win count = 105 Win ratio = 37.234042553191486
Epoch 282: err = 0.10008768737316132: Win count = 106 Win ratio = 37.455830388692576
Epoch 283: err = 0.07775689335539937: Win count = 107 Win ratio = 37.67605633802817
Epoch 284: err = 0.11389708565548062: Win count = 108 Win ratio = 37.89473684210527
Epoch 285: err = 0.09888859931379557: Win count = 109 Win ratio = 38.111888111888106
Epoch 286: err = 0.10212073288857937: Win count = 110 Win ratio = 38.32752613240418
Epoch 287: err = 0.14180093072354794: Win count = 111 Win ratio = 38.54166666666667
Epoch 288: err = 0.09345205128192902: Win count = 112 Win ratio = 38.75432525951557
Epoch 289: err = 0.053975182585418224: Win count = 112 Win ratio = 38.62068965517241
Epoch 290: err = 0.08458134019747376: Win count = 113 Win ratio = 38.83161512027492
Epoch 291: err = 0.09283559676259756: Win count = 114 Win ratio = 39.04109589041096
Epoch 292: err = 0.0734427603892982: Win count = 115 Win ratio = 39.249146757679185
Epoch 293: err = 0.09519645664840937: Win count = 116 Win ratio = 39.455782312925166
Epoch 294: err = 0.08602534281089902: Win count = 117 Win ratio = 39.66101694915255
Epoch 295: err = 0.10513025335967541: Win count = 118 Win ratio = 39.86486486486486
Epoch 296: err = 0.08710489561781287: Win count = 119 Win ratio = 40.06734006734007
Epoch 297: err = 0.09991425415500998: Win count = 120 Win ratio = 40.26845637583892
Epoch 298: err = 0.10896686278283596: Win count = 121 Win ratio = 40.468227424749166

Epoch 299: err = 0.05817233677953482: Win count = 122 Win ratio = 40.666666666666664
Epoch 300: err = 0.06714716833084822: Win count = 123 Win ratio = 40.863787375415285
Epoch 301: err = 0.11066560028120875: Win count = 123 Win ratio = 40.728476821192054
Epoch 302: err = 0.07632107893005013: Win count = 124 Win ratio = 40.92409240924093
Epoch 303: err = 0.07045093830674887: Win count = 125 Win ratio = 41.118421052631575
Epoch 304: err = 0.10706557240337133: Win count = 126 Win ratio = 41.31147540983607
Epoch 305: err = 0.07666993979364634: Win count = 127 Win ratio = 41.50326797385621
Epoch 306: err = 0.07974411733448505: Win count = 128 Win ratio = 41.69381107491857
Epoch 307: err = 0.05425239587202668: Win count = 129 Win ratio = 41.883116883116884
Epoch 308: err = 0.06228772131726146: Win count = 130 Win ratio = 42.07119741100323
Epoch 309: err = 0.06191363325342536: Win count = 131 Win ratio = 42.25806451612903
Epoch 310: err = 0.05079396814107895: Win count = 132 Win ratio = 42.443729903536976
Epoch 311: err = 0.05655255145393312: Win count = 133 Win ratio = 42.628205128205124
Epoch 312: err = 0.0512783951126039: Win count = 134 Win ratio = 42.81150159744409
Epoch 313: err = 0.08246976998634636: Win count = 135 Win ratio = 42.99363057324841
Epoch 314: err = 0.06801556353457272: Win count = 136 Win ratio = 43.17460317460318
Epoch 315: err = 0.06591094098985195: Win count = 137 Win ratio = 43.35443037974683
Epoch 316: err = 0.07753805909305811: Win count = 138 Win ratio = 43.53312302839117
Epoch 317: err = 0.07643223414197564: Win count = 138 Win ratio = 43.39622641509434
Epoch 318: err = 0.06639134231954813: Win count = 139 Win ratio = 43.573667711598745
Epoch 319: err = 0.040663027903065085: Win count = 140 Win ratio = 43.75
Epoch 320: err = 0.04255857993848622: Win count = 141 Win ratio = 43.925233644859816
Epoch 321: err = 0.02801830694079399: Win count = 142 Win ratio = 44.099378881987576
Epoch 322: err = 0.026257609017193317: Win count = 143 Win ratio = 44.27244582043344
Epoch 323: err = 0.03568953764624894: Win count = 144 Win ratio = 44.444444444444444
Epoch 324: err = 0.03766967845149338: Win count = 144 Win ratio = 44.30769230769231
Epoch 325: err = 0.03095076442696154: Win count = 145 Win ratio = 44.47852760736196
Epoch 326: err = 0.03816063818521798: Win count = 146 Win ratio = 44.64831804281346
Epoch 327: err = 0.03622238524258137: Win count = 147 Win ratio = 44.81707317073171
Epoch 328: err = 0.03260565036907792: Win count = 147 Win ratio = 44.680851063829785
Epoch 329: err = 0.03089148853905499: Win count = 148 Win ratio = 44.84848484848485
Epoch 330: err = 0.03446335601620376: Win count = 149 Win ratio = 45.01510574018127
Epoch 331: err = 0.026642409153282642: Win count = 150 Win ratio = 45.18072289156627
Epoch 332: err = 0.027539491653442383: Win count = 151 Win ratio = 45.34534534534534
Epoch 333: err = 0.035845580510795116: Win count = 152 Win ratio = 45.50898203592814
Epoch 334: err = 0.03309162543155253: Win count = 153 Win ratio = 45.67164179104478
Epoch 335: err = 0.026661335490643978: Win count = 154 Win ratio = 45.83333333333333
Epoch 336: err = 0.02563447947613895: Win count = 155 Win ratio = 45.99406528189911
Epoch 337: err = 0.025665166787803173: Win count = 156 Win ratio = 46.15384615384615
Epoch 338: err = 0.02361704525537789: Win count = 157 Win ratio = 46.312684365781706
Epoch 339: err = 0.02213092742022127: Win count = 157 Win ratio = 46.1764705882353
Epoch 340: err = 0.018194570671766996: Win count = 158 Win ratio = 46.33431085043988
Epoch 341: err = 0.02135899057611823: Win count = 158 Win ratio = 46.198830409356724
Epoch 342: err = 0.03525615367107093: Win count = 158 Win ratio = 46.06413994169096
Epoch 343: err = 0.029012358747422695: Win count = 159 Win ratio = 46.22093023255814
Epoch 344: err = 0.01665875338949263: Win count = 160 Win ratio = 46.3768115942029
Epoch 345: err = 0.015602928237058222: Win count = 161 Win ratio = 46.53179190751445
Epoch 346: err = 0.015137258335016668: Win count = 162 Win ratio = 46.68587896253602
Epoch 347: err = 0.03032277710735798: Win count = 163 Win ratio = 46.839080459770116
Epoch 348: err = 0.027999445097520947: Win count = 164 Win ratio = 46.99140401146132
Epoch 349: err = 0.03197944036219269: Win count = 165 Win ratio = 47.14285714285714
Epoch 350: err = 0.015815933933481574: Win count = 166 Win ratio = 47.293447293447294
Epoch 351: err = 0.02986084611620754: Win count = 167 Win ratio = 47.44318181818182
Epoch 352: err = 0.022900561685673892: Win count = 168 Win ratio = 47.59206798866856
Epoch 353: err = 0.018176296027377248: Win count = 169 Win ratio = 47.740112994350284
Epoch 354: err = 0.020587756705936044: Win count = 170 Win ratio = 47.88732394366197
Epoch 355: err = 0.022892315697390586: Win count = 171 Win ratio = 48.03370786516854
Epoch 356: err = 0.02858372952323407: Win count = 172 Win ratio = 48.17927170868347
Epoch 357: err = 0.020015172311104834: Win count = 173 Win ratio = 48.324022346368714
Epoch 358: err = 0.01753783581079915: Win count = 174 Win ratio = 48.467966573816156

Epoch 359: err = 0.019201011396944523: Win count = 174 Win ratio = 48.333333333333336
Epoch 360: err = 0.016015480738133192: Win count = 175 Win ratio = 48.476454293628805
Epoch 361: err = 0.017259390326216817: Win count = 176 Win ratio = 48.61878453038674
Epoch 362: err = 0.01426272711250931: Win count = 177 Win ratio = 48.760330578512395
Epoch 363: err = 0.018214279087260365: Win count = 177 Win ratio = 48.62637362637363
Epoch 364: err = 0.014217701042070985: Win count = 177 Win ratio = 48.49315068493151
Epoch 365: err = 0.01176273263990879: Win count = 178 Win ratio = 48.63387978142077
Epoch 366: err = 0.013543022214435041: Win count = 179 Win ratio = 48.77384196185286
Epoch 367: err = 0.01761997409630567: Win count = 180 Win ratio = 48.91304347826087
Epoch 368: err = 0.013584851403720677: Win count = 181 Win ratio = 49.05149051490515
Epoch 369: err = 0.010596422420348972: Win count = 182 Win ratio = 49.18918918918919
Epoch 370: err = 0.009797979204449803: Win count = 182 Win ratio = 49.056603773584904
Epoch 371: err = 0.01213558774907142: Win count = 183 Win ratio = 49.193548387096776
Epoch 372: err = 0.01165875163860619: Win count = 183 Win ratio = 49.06166219839142
Epoch 373: err = 0.013902777864132077: Win count = 183 Win ratio = 48.93048128342246
Epoch 374: err = 0.03714465972734615: Win count = 184 Win ratio = 49.06666666666666
Epoch 375: err = 0.021518958848901093: Win count = 185 Win ratio = 49.202127659574465
Epoch 376: err = 0.018055692547932267: Win count = 186 Win ratio = 49.3368700265252
Epoch 377: err = 0.015332438750192523: Win count = 187 Win ratio = 49.47089947089947
Epoch 378: err = 0.017871597432531416: Win count = 188 Win ratio = 49.6042216358839
Epoch 379: err = 0.009892682021018118: Win count = 189 Win ratio = 49.73684210526316
Epoch 380: err = 0.01947609888156876: Win count = 190 Win ratio = 49.86876640419948
Epoch 381: err = 0.017772015184164047: Win count = 191 Win ratio = 50.0
Epoch 382: err = 0.02283422090113163: Win count = 192 Win ratio = 50.13054830287206
Epoch 383: err = 0.013126179110258818: Win count = 193 Win ratio = 50.260416666666664
Epoch 384: err = 0.024382224888540804: Win count = 194 Win ratio = 50.38961038961038
Epoch 385: err = 0.011332634196151048: Win count = 195 Win ratio = 50.51813471502591
Epoch 386: err = 0.015108467428945005: Win count = 196 Win ratio = 50.64599483204134
Epoch 387: err = 0.012438031437341124: Win count = 197 Win ratio = 50.77319587628865
Epoch 388: err = 0.011754794919397682: Win count = 198 Win ratio = 50.899742930591266
Epoch 389: err = 0.00802934862440452: Win count = 199 Win ratio = 51.02564102564102
Epoch 390: err = 0.014402099186554551: Win count = 199 Win ratio = 50.89514066496164
Epoch 391: err = 0.01185251446440816: Win count = 200 Win ratio = 51.02040816326531
Epoch 392: err = 0.06141596962697804: Win count = 201 Win ratio = 51.14503816793893
Epoch 393: err = 0.048762253718450665: Win count = 202 Win ratio = 51.26903553299492
Epoch 394: err = 0.024148533586412668: Win count = 203 Win ratio = 51.392405063291136
Epoch 395: err = 0.01426446542609483: Win count = 204 Win ratio = 51.515151515151516
Epoch 396: err = 0.008576535095926374: Win count = 205 Win ratio = 51.63727959697732
Epoch 397: err = 0.025346130598336458: Win count = 206 Win ratio = 51.75879396984925
Epoch 398: err = 0.03312408260535449: Win count = 207 Win ratio = 51.8796992481203
Epoch 399: err = 0.024572775524575263: Win count = 208 Win ratio = 52.0
Epoch 400: err = 0.028751338366419077: Win count = 209 Win ratio = 52.119700748129674
Epoch 401: err = 0.014407808834221214: Win count = 210 Win ratio = 52.23880597014925
Epoch 402: err = 0.01796950149582699: Win count = 211 Win ratio = 52.357320099255574
Epoch 403: err = 0.02251176704885438: Win count = 212 Win ratio = 52.475247524752476
Epoch 404: err = 0.009385086013935506: Win count = 213 Win ratio = 52.59259259259259
Epoch 405: err = 0.008362300315639004: Win count = 214 Win ratio = 52.70935960591133
Epoch 406: err = 0.02717274078167975: Win count = 215 Win ratio = 52.825552825552826
Epoch 407: err = 0.022773829521611333: Win count = 216 Win ratio = 52.94117647058824
Epoch 408: err = 0.015608081128448248: Win count = 216 Win ratio = 52.81173594132029
Epoch 409: err = 0.015782252710778266: Win count = 217 Win ratio = 52.926829268292686
Epoch 410: err = 0.011688172642607242: Win count = 218 Win ratio = 53.04136253041363
Epoch 411: err = 0.013152395724318922: Win count = 219 Win ratio = 53.15533980582524
Epoch 412: err = 0.017532882280647755: Win count = 220 Win ratio = 53.26876513317191
Epoch 413: err = 0.02050428540678695: Win count = 221 Win ratio = 53.3816425120773
Epoch 414: err = 0.027175483002793044: Win count = 222 Win ratio = 53.493975903614455
Epoch 415: err = 0.022306436439976096: Win count = 223 Win ratio = 53.605769230769226
Epoch 416: err = 0.009561945626046509: Win count = 224 Win ratio = 53.71702637889688
Epoch 417: err = 0.030605439795181155: Win count = 225 Win ratio = 53.827751196172244
Epoch 418: err = 0.017034815449733287: Win count = 226 Win ratio = 53.93794749403341

Epoch 419: err = 0.021667772554792464: Win count = 227 Win ratio = 54.047619047619044
Epoch 420: err = 0.022115676081739366: Win count = 228 Win ratio = 54.156769596199524
Epoch 421: err = 0.015964234014973044: Win count = 229 Win ratio = 54.2654028436019
Epoch 422: err = 0.012758478726027533: Win count = 230 Win ratio = 54.37352245862884
Epoch 423: err = 0.013666419545188546: Win count = 231 Win ratio = 54.48113207547169
Epoch 424: err = 0.01354261819506064: Win count = 232 Win ratio = 54.58823529411765
Epoch 425: err = 0.013646140811033547: Win count = 233 Win ratio = 54.694835680751176
Epoch 426: err = 0.009939704672433436: Win count = 233 Win ratio = 54.56674473067916
Epoch 427: err = 0.012050110613927245: Win count = 234 Win ratio = 54.67289719626168
Epoch 428: err = 0.016932452330365777: Win count = 235 Win ratio = 54.77855477855478
Epoch 429: err = 0.010486344050150365: Win count = 236 Win ratio = 54.88372093023256
Epoch 430: err = 0.008834070933517069: Win count = 237 Win ratio = 54.988399071925755
Epoch 431: err = 0.013825631816871464: Win count = 238 Win ratio = 55.092592592592595
Epoch 432: err = 0.018514614377636462: Win count = 238 Win ratio = 54.96535796766744
Epoch 433: err = 0.015031907358206809: Win count = 239 Win ratio = 55.06912442396313
Epoch 434: err = 0.010538582631852478: Win count = 240 Win ratio = 55.172413793103445
Epoch 435: err = 0.012475281953811646: Win count = 241 Win ratio = 55.27522935779816
Epoch 436: err = 0.008695928903762251: Win count = 242 Win ratio = 55.37757437070938
Epoch 437: err = 0.010464628459885716: Win count = 243 Win ratio = 55.47945205479452
Epoch 438: err = 0.007947085599880666: Win count = 243 Win ratio = 55.35307517084283
Epoch 439: err = 0.011698611080646515: Win count = 244 Win ratio = 55.45454545454545
Epoch 440: err = 0.006549458979861811: Win count = 245 Win ratio = 55.55555555555556
Epoch 441: err = 0.006934285745956004: Win count = 245 Win ratio = 55.42986425339367
Epoch 442: err = 0.020934739091899246: Win count = 246 Win ratio = 55.53047404063205
Epoch 443: err = 0.007628364721313119: Win count = 247 Win ratio = 55.63063063063063
Epoch 444: err = 0.009951301792170852: Win count = 248 Win ratio = 55.73033707865168
Epoch 445: err = 0.007892639725469053: Win count = 249 Win ratio = 55.82959641255605
Epoch 446: err = 0.004404805920785293: Win count = 250 Win ratio = 55.92841163310962
Epoch 447: err = 0.00669068782008253: Win count = 251 Win ratio = 56.02678571428571
Epoch 448: err = 0.003939114132663235: Win count = 252 Win ratio = 56.12472160356348
Epoch 449: err = 0.005102792492834851: Win count = 253 Win ratio = 56.222222222222214
Epoch 450: err = 0.007050473737763241: Win count = 254 Win ratio = 56.31929046563193
Epoch 451: err = 0.002959817385999486: Win count = 255 Win ratio = 56.415929203539825
Epoch 452: err = 0.0035931644670199603: Win count = 255 Win ratio = 56.29139072847682
Epoch 453: err = 0.0067391403135843575: Win count = 256 Win ratio = 56.38766519823789
Epoch 454: err = 0.008670789131429046: Win count = 257 Win ratio = 56.48351648351648
Epoch 455: err = 0.017548715346492827: Win count = 258 Win ratio = 56.57894736842105
Epoch 456: err = 0.02154904027702287: Win count = 259 Win ratio = 56.67396061269147
Epoch 457: err = 0.013655219343490899: Win count = 260 Win ratio = 56.76855895196506
Epoch 458: err = 0.01410910757840611: Win count = 261 Win ratio = 56.86274509803921
Epoch 459: err = 0.012983892112970352: Win count = 261 Win ratio = 56.739130434782616
Epoch 460: err = 0.007120988389942795: Win count = 261 Win ratio = 56.61605206073753
Epoch 461: err = 0.0035070487792836502: Win count = 262 Win ratio = 56.709956709956714
Epoch 462: err = 0.005650125356623903: Win count = 263 Win ratio = 56.803455723542115
Epoch 463: err = 0.0031635896302759647: Win count = 264 Win ratio = 56.896551724137936
Epoch 464: err = 0.006297581261605956: Win count = 265 Win ratio = 56.98924731182796
Epoch 465: err = 0.007430131852743216: Win count = 266 Win ratio = 57.08154506437768
Epoch 466: err = 0.004541232672636397: Win count = 267 Win ratio = 57.17344753747323
Epoch 467: err = 0.0025605571718188003: Win count = 268 Win ratio = 57.26495726495726
Epoch 468: err = 0.0023327597591560334: Win count = 268 Win ratio = 57.14285714285714
Epoch 469: err = 0.005481181317009032: Win count = 269 Win ratio = 57.234042553191486
Epoch 470: err = 0.003628199774539098: Win count = 270 Win ratio = 57.324840764331206
Epoch 471: err = 0.00531049893470481: Win count = 271 Win ratio = 57.41525423728814
Epoch 472: err = 0.0017608805210329592: Win count = 272 Win ratio = 57.505285412262154
Epoch 473: err = 0.0015175891676335596: Win count = 272 Win ratio = 57.383966244725734
Epoch 474: err = 0.0021991666144458577: Win count = 273 Win ratio = 57.47368421052632
Epoch 475: err = 0.009760296670719981: Win count = 274 Win ratio = 57.56302521008403
Epoch 476: err = 0.0031114992016227916: Win count = 275 Win ratio = 57.65199161425576
Epoch 477: err = 0.004728120475192554: Win count = 276 Win ratio = 57.74058577405857
Epoch 478: err = 0.006485084901214577: Win count = 277 Win ratio = 57.82881002087683

Epoch 479: err = 0.004929268572595902: Win count = 278 Win ratio = 57.91666666666667
Epoch 480: err = 0.0028955097659491003: Win count = 279 Win ratio = 58.004158004158
Epoch 481: err = 0.0036908581823809072: Win count = 280 Win ratio = 58.09128630705395
Epoch 482: err = 0.0050354157137917355: Win count = 280 Win ratio = 57.971014492753625
Epoch 483: err = 0.006602250650757924: Win count = 281 Win ratio = 58.057851239669425
Epoch 484: err = 0.002991471308632754: Win count = 281 Win ratio = 57.938144329896915
Epoch 485: err = 0.0024535736010875553: Win count = 282 Win ratio = 58.0246913580247
Epoch 486: err = 0.002170934691093862: Win count = 283 Win ratio = 58.110882956878854
Epoch 487: err = 0.001611452316865325: Win count = 284 Win ratio = 58.19672131147541
Epoch 488: err = 0.001471861389291007: Win count = 285 Win ratio = 58.282208588957054
Epoch 489: err = 0.0014972374628996477: Win count = 286 Win ratio = 58.36734693877551
Epoch 490: err = 0.001993024969124235: Win count = 287 Win ratio = 58.45213849287168
Epoch 491: err = 0.0020836329349549487: Win count = 288 Win ratio = 58.536585365853654
Epoch 492: err = 0.0021761338502983563: Win count = 289 Win ratio = 58.620689655172406
Epoch 493: err = 0.0021067147608846426: Win count = 290 Win ratio = 58.70445344129555
Epoch 494: err = 0.0020389408673509024: Win count = 291 Win ratio = 58.78787878787879
Epoch 495: err = 0.0012950265518156812: Win count = 292 Win ratio = 58.87096774193549
Epoch 496: err = 0.001200712358695455: Win count = 293 Win ratio = 58.95372233400402
Epoch 497: err = 0.001171796306152828: Win count = 294 Win ratio = 59.036144578313255
Epoch 498: err = 0.0017188020501635037: Win count = 295 Win ratio = 59.118236472945895
Epoch 499: err = 0.002911483148636762: Win count = 296 Win ratio = 59.199999999999996
Epoch 500: err = 0.0023948850503074937: Win count = 297 Win ratio = 59.2814371257485
Epoch 501: err = 0.0018145096764783375: Win count = 298 Win ratio = 59.36254980079681
Epoch 502: err = 0.0013857138328603469: Win count = 298 Win ratio = 59.24453280318092
Epoch 503: err = 0.0021760021554655395: Win count = 299 Win ratio = 59.32539682539682
Epoch 504: err = 0.001949019031599164: Win count = 300 Win ratio = 59.4059405940594
Epoch 505: err = 0.0013624212224385701: Win count = 301 Win ratio = 59.48616600790514
Epoch 506: err = 0.0012734688425553031: Win count = 302 Win ratio = 59.56607495069034
Epoch 507: err = 0.0015045141335576773: Win count = 303 Win ratio = 59.645669291338585
Epoch 508: err = 0.002900242143368814: Win count = 303 Win ratio = 59.52848722986247
Epoch 509: err = 0.0017496722939540632: Win count = 304 Win ratio = 59.6078431372549
Epoch 510: err = 0.0020361694841994904: Win count = 305 Win ratio = 59.68688845401174
Epoch 511: err = 0.0025919596264429856: Win count = 306 Win ratio = 59.765625
Epoch 512: err = 0.0012905814655823633: Win count = 307 Win ratio = 59.84405458089669
Epoch 513: err = 0.001276167509786319: Win count = 308 Win ratio = 59.92217898832685
Epoch 514: err = 0.001380928690196015: Win count = 309 Win ratio = 60.0
Epoch 515: err = 0.0015700018557254225: Win count = 310 Win ratio = 60.07751937984496
Epoch 516: err = 0.0011317579337628558: Win count = 310 Win ratio = 59.96131528046421
Epoch 517: err = 0.0008475058675685432: Win count = 311 Win ratio = 60.03861003861004
Epoch 518: err = 0.0007666381425224245: Win count = 312 Win ratio = 60.115606936416185
Epoch 519: err = 0.0009597324169590138: Win count = 313 Win ratio = 60.192307692307686
Epoch 520: err = 0.0009685657532827463: Win count = 314 Win ratio = 60.26871401151631
Epoch 521: err = 0.0009872634909697808: Win count = 315 Win ratio = 60.3448275862069
Epoch 522: err = 0.0007423817514791153: Win count = 316 Win ratio = 60.42065009560229
Epoch 523: err = 0.001351694496406708: Win count = 317 Win ratio = 60.49618320610687
Epoch 524: err = 0.0012167180066171568: Win count = 318 Win ratio = 60.57142857142858
Epoch 525: err = 0.0020657926725107245: Win count = 319 Win ratio = 60.646387832699624
Epoch 526: err = 0.00208246612601215: Win count = 320 Win ratio = 60.72106261859582
Epoch 527: err = 0.0018315884517505765: Win count = 321 Win ratio = 60.79545454545454
Epoch 528: err = 0.0011236271311645396: Win count = 322 Win ratio = 60.86956521739131
Epoch 529: err = 0.000904583645024104: Win count = 323 Win ratio = 60.94339622641509
Epoch 530: err = 0.0010489301748748403: Win count = 324 Win ratio = 61.016949152542374
Epoch 531: err = 0.002136725299351383: Win count = 325 Win ratio = 61.09022556390977
Epoch 532: err = 0.0017123850993812084: Win count = 326 Win ratio = 61.16322701688556
Epoch 533: err = 0.0012212801339046564: Win count = 327 Win ratio = 61.23595505617978
Epoch 534: err = 0.0013421322946669534: Win count = 328 Win ratio = 61.308411214953274
Epoch 535: err = 0.0012641119246836752: Win count = 329 Win ratio = 61.38059701492538
Epoch 536: err = 0.0011981196876149625: Win count = 330 Win ratio = 61.452513966480446
Epoch 537: err = 0.0013092980152578093: Win count = 331 Win ratio = 61.52416356877324
Epoch 538: err = 0.0010485699676792137: Win count = 332 Win ratio = 61.59554730983302

Epoch 539: err = 0.0007537848032370675: Win count = 333 Win ratio = 61.66666666666667
Epoch 540: err = 0.0006080316270526964: Win count = 334 Win ratio = 61.737523105360445
Epoch 541: err = 0.001369692512525944: Win count = 335 Win ratio = 61.80811808118081
Epoch 542: err = 0.0009141394657490309: Win count = 336 Win ratio = 61.87845303867403
Epoch 543: err = 0.0009634218658902682: Win count = 337 Win ratio = 61.94852941176471
Epoch 544: err = 0.0011540660561877303: Win count = 338 Win ratio = 62.018348623853214
Epoch 545: err = 0.0010287858058291022: Win count = 338 Win ratio = 61.904761904761905
Epoch 546: err = 0.0006543605668412056: Win count = 339 Win ratio = 61.97440585009141
Epoch 547: err = 0.0007187452174548525: Win count = 340 Win ratio = 62.04379562043796
Epoch 548: err = 0.0005858156837348361: Win count = 341 Win ratio = 62.11293260473588
Epoch 549: err = 0.001649932692089351: Win count = 342 Win ratio = 62.18181818181818
Epoch 550: err = 0.0032891427326831035: Win count = 343 Win ratio = 62.25045372050817
Epoch 551: err = 0.0012648502379306592: Win count = 344 Win ratio = 62.31884057971014
Epoch 552: err = 0.0013831290561938658: Win count = 345 Win ratio = 62.38698010849909
Epoch 553: err = 0.0005629918196063954: Win count = 346 Win ratio = 62.454873646209386
Epoch 554: err = 0.0008061909502430353: Win count = 347 Win ratio = 62.52252252252253
Epoch 555: err = 0.0016134845864144154: Win count = 348 Win ratio = 62.589928057553955
Epoch 556: err = 0.0007590194290969521: Win count = 349 Win ratio = 62.65709156193896
Epoch 557: err = 0.0009970637474907562: Win count = 350 Win ratio = 62.72401433691756
Epoch 558: err = 0.0017634314426686615: Win count = 351 Win ratio = 62.7906976744186
Epoch 559: err = 0.0013860059407306835: Win count = 352 Win ratio = 62.857142857142854
Epoch 560: err = 0.0011009589397872332: Win count = 353 Win ratio = 62.92335115864528
Epoch 561: err = 0.00116368658200372: Win count = 354 Win ratio = 62.989323843416365
Epoch 562: err = 0.0017336561795673333: Win count = 355 Win ratio = 63.05506216696269
Epoch 563: err = 0.001128789910580963: Win count = 356 Win ratio = 63.12056737588653
Epoch 564: err = 0.0017862659369711764: Win count = 357 Win ratio = 63.1858407079646
Epoch 565: err = 0.001543204220070038: Win count = 358 Win ratio = 63.25088339222616
Epoch 566: err = 0.0009855587195488624: Win count = 359 Win ratio = 63.31569664902999
Epoch 567: err = 0.0009665842735557817: Win count = 360 Win ratio = 63.38028169014085
Epoch 568: err = 0.0011959669063799083: Win count = 361 Win ratio = 63.44463971880492
Epoch 569: err = 0.0007207714406831656: Win count = 362 Win ratio = 63.50877192982456
Epoch 570: err = 0.0014029652338649612: Win count = 363 Win ratio = 63.57267950963222
Epoch 571: err = 0.0011480477805889677: Win count = 364 Win ratio = 63.63636363636363
Epoch 572: err = 0.0008079092840489466: Win count = 365 Win ratio = 63.6998254799302
Epoch 573: err = 0.0010683940163289662: Win count = 366 Win ratio = 63.76306620209059
Epoch 574: err = 0.0006941135216038674: Win count = 367 Win ratio = 63.82608695652174
Epoch 575: err = 0.0006533111081807874: Win count = 368 Win ratio = 63.888888888888886
Epoch 576: err = 0.0007470078380720224: Win count = 369 Win ratio = 63.951473136915084
Epoch 577: err = 0.0005988584744045511: Win count = 370 Win ratio = 64.01384083044984
Epoch 578: err = 0.0007717337830399629: Win count = 371 Win ratio = 64.07599309153713
Epoch 579: err = 0.0007250569324241951: Win count = 371 Win ratio = 63.96551724137931
Epoch 580: err = 0.0004004921702289721: Win count = 372 Win ratio = 64.02753872633392
Epoch 581: err = 0.000705151727743214: Win count = 373 Win ratio = 64.08934707903781
Epoch 582: err = 0.0005883879166503903: Win count = 374 Win ratio = 64.15094339622641
Epoch 583: err = 0.0008232318814407336: Win count = 375 Win ratio = 64.21232876712328
Epoch 584: err = 0.0006873770289530512: Win count = 376 Win ratio = 64.27350427350427
Epoch 585: err = 0.0007441806355927838: Win count = 377 Win ratio = 64.3344709897611
Epoch 586: err = 0.000731235253624618: Win count = 378 Win ratio = 64.39522998296422
Epoch 587: err = 0.0006567297841684194: Win count = 379 Win ratio = 64.45578231292517
Epoch 588: err = 0.0007797277758072596: Win count = 380 Win ratio = 64.51612903225806
Epoch 589: err = 0.0006411073591152672: Win count = 381 Win ratio = 64.57627118644068
Epoch 590: err = 0.00077411567690433: Win count = 382 Win ratio = 64.63620981387479
Epoch 591: err = 0.0005498289028764702: Win count = 383 Win ratio = 64.69594594594594
Epoch 592: err = 0.0004124473944102647: Win count = 384 Win ratio = 64.75548060708263
Epoch 593: err = 0.0006557687338499818: Win count = 385 Win ratio = 64.81481481481481
Epoch 594: err = 0.0005893746347283013: Win count = 386 Win ratio = 64.87394957983193
Epoch 595: err = 0.0011479420572868548: Win count = 387 Win ratio = 64.93288590604027
Epoch 596: err = 0.000594989560340764: Win count = 388 Win ratio = 64.99162479061977
Epoch 597: err = 0.0008897867992345709: Win count = 389 Win ratio = 65.05016722408027
Epoch 598: err = 0.001039391327140038: Win count = 390 Win ratio = 65.1085141903172

Epoch 599: err = 0.0006647768677794375: Win count = 391 Win ratio = 65.16666666666666
Epoch 600: err = 0.0021510075184778543: Win count = 392 Win ratio = 65.22462562396007
Epoch 601: err = 0.0018337255678488873: Win count = 393 Win ratio = 65.28239202657808
Epoch 602: err = 0.0014407731359824538: Win count = 394 Win ratio = 65.33996683250415
Epoch 603: err = 0.001390943769365549: Win count = 395 Win ratio = 65.39735099337747
Epoch 604: err = 0.0024416425367235206: Win count = 396 Win ratio = 65.45454545454545
Epoch 605: err = 0.001946297867107205: Win count = 397 Win ratio = 65.51155115511551
Epoch 606: err = 0.002910954790422693: Win count = 398 Win ratio = 65.56836902800659
Epoch 607: err = 0.00223253318836214: Win count = 399 Win ratio = 65.625
Epoch 608: err = 0.001806453787139617: Win count = 400 Win ratio = 65.68144499178982
Epoch 609: err = 0.001888536658952944: Win count = 401 Win ratio = 65.73770491803279
Epoch 610: err = 0.0010916916980932: Win count = 402 Win ratio = 65.79378068739771
Epoch 611: err = 0.005167993105715141: Win count = 403 Win ratio = 65.84967320261438
Epoch 612: err = 0.014937203217414208: Win count = 404 Win ratio = 65.90538336052202
Epoch 613: err = 0.002636298020661343: Win count = 405 Win ratio = 65.96091205211727
Epoch 614: err = 0.0040453216643072665: Win count = 406 Win ratio = 66.01626016260163
Epoch 615: err = 0.0028640452583204024: Win count = 407 Win ratio = 66.07142857142857
Epoch 616: err = 0.0018065862386720255: Win count = 407 Win ratio = 65.96434359805511
Epoch 617: err = 0.0016062902723206207: Win count = 408 Win ratio = 66.01941747572816
Epoch 618: err = 0.0014049930105102248: Win count = 409 Win ratio = 66.07431340872375
Epoch 619: err = 0.003501906459860038: Win count = 410 Win ratio = 66.12903225806451
Epoch 620: err = 0.0018843987782020122: Win count = 411 Win ratio = 66.18357487922705
Epoch 621: err = 0.0024621300253784284: Win count = 412 Win ratio = 66.2379421221865
Epoch 622: err = 0.0013408251616056077: Win count = 413 Win ratio = 66.29213483146067
Epoch 623: err = 0.0016083787777461112: Win count = 414 Win ratio = 66.34615384615384
Epoch 624: err = 0.001985984454222489: Win count = 415 Win ratio = 66.4
Epoch 625: err = 0.0031004595948616043: Win count = 416 Win ratio = 66.45367412140575
Epoch 626: err = 0.0026276213029632345: Win count = 417 Win ratio = 66.50717703349282
Epoch 627: err = 0.002727264254644979: Win count = 418 Win ratio = 66.56050955414013
Epoch 628: err = 0.0015917674827505834: Win count = 419 Win ratio = 66.61367249602543
Epoch 629: err = 0.0015342551305366214: Win count = 420 Win ratio = 66.66666666666666
Epoch 630: err = 0.002688314314582385: Win count = 421 Win ratio = 66.71949286846277
Epoch 631: err = 0.0015305382185033523: Win count = 422 Win ratio = 66.77215189873418
Epoch 632: err = 0.0025271010017604567: Win count = 423 Win ratio = 66.82464454976304
Epoch 633: err = 0.0021138841766514815: Win count = 424 Win ratio = 66.8769716088328
Epoch 634: err = 0.0013413802444119938: Win count = 425 Win ratio = 66.92913385826772
Epoch 635: err = 0.0015770665122545324: Win count = 426 Win ratio = 66.98113207547169
Epoch 636: err = 0.00149352216249099: Win count = 426 Win ratio = 66.87598116169545
Epoch 637: err = 0.0018579159659566358: Win count = 427 Win ratio = 66.92789968652038
Epoch 638: err = 0.0009564784559188411: Win count = 427 Win ratio = 66.82316118935837
Epoch 639: err = 0.002230796526418999: Win count = 428 Win ratio = 66.875
Epoch 640: err = 0.0016364816274290206: Win count = 429 Win ratio = 66.92667706708268
Epoch 641: err = 0.0010633537058311049: Win count = 429 Win ratio = 66.82242990654206
Epoch 642: err = 0.0014162240986479446: Win count = 430 Win ratio = 66.87402799377917
Epoch 643: err = 0.0013655881266458891: Win count = 431 Win ratio = 66.92546583850931
Epoch 644: err = 0.0018881110299844295: Win count = 432 Win ratio = 66.97674418604652
Epoch 645: err = 0.0015296210694941692: Win count = 433 Win ratio = 67.02786377708978
Epoch 646: err = 0.0015730249797343276: Win count = 434 Win ratio = 67.07882534775888
Epoch 647: err = 0.0013024425861658528: Win count = 435 Win ratio = 67.12962962962963
Epoch 648: err = 0.0011582649603951722: Win count = 436 Win ratio = 67.18027734976889
Epoch 649: err = 0.0008357769311260199: Win count = 437 Win ratio = 67.23076923076923
Epoch 650: err = 0.001159263847512193: Win count = 438 Win ratio = 67.2811059907834
Epoch 651: err = 0.0017270629323320463: Win count = 439 Win ratio = 67.33128834355828
Epoch 652: err = 0.0012972908225492574: Win count = 440 Win ratio = 67.3813169984686
Epoch 653: err = 0.0015295493649318814: Win count = 441 Win ratio = 67.43119266055045
Epoch 654: err = 0.001288116000688396: Win count = 442 Win ratio = 67.48091603053436
Epoch 655: err = 0.0009228586532117333: Win count = 443 Win ratio = 67.53048780487805
Epoch 656: err = 0.0011913895978068467: Win count = 444 Win ratio = 67.57990867579909
Epoch 657: err = 0.0010147481480089482: Win count = 445 Win ratio = 67.62917933130699
Epoch 658: err = 0.0008298034699691925: Win count = 446 Win ratio = 67.6783004552352

Epoch 659: err = 0.0011410954175516963: Win count = 447 Win ratio = 67.72727272727272
Epoch 660: err = 0.0008614328944531735: Win count = 448 Win ratio = 67.77609682299546
Epoch 661: err = 0.0009202990077028517: Win count = 449 Win ratio = 67.82477341389728
Epoch 662: err = 0.001236085117852781: Win count = 450 Win ratio = 67.87330316742081
Epoch 663: err = 0.0012130141840316355: Win count = 451 Win ratio = 67.92168674698796
Epoch 664: err = 0.0011077563103754073: Win count = 452 Win ratio = 67.96992481203009
Epoch 665: err = 0.0009455029994569486: Win count = 453 Win ratio = 68.01801801801803
Epoch 666: err = 0.0009139518624579068: Win count = 454 Win ratio = 68.06596701649175
Epoch 667: err = 0.0008007105225260602: Win count = 455 Win ratio = 68.11377245508982
Epoch 668: err = 0.0007996883505256847: Win count = 456 Win ratio = 68.16143497757847
Epoch 669: err = 0.0007703761802986264: Win count = 456 Win ratio = 68.05970149253731
Epoch 670: err = 0.001266395760467276: Win count = 457 Win ratio = 68.10730253353205
Epoch 671: err = 0.0009186759343720041: Win count = 458 Win ratio = 68.15476190476191
Epoch 672: err = 0.0012545877398224548: Win count = 459 Win ratio = 68.20208023774146
Epoch 673: err = 0.0008167330888682045: Win count = 460 Win ratio = 68.24925816023739
Epoch 674: err = 0.0006135380535852164: Win count = 461 Win ratio = 68.2962962962963
Epoch 675: err = 0.0008833780593704432: Win count = 462 Win ratio = 68.34319526627219
Epoch 676: err = 0.0006187957151269075: Win count = 463 Win ratio = 68.38995568685377
Epoch 677: err = 0.0008989002017187886: Win count = 464 Win ratio = 68.43657817109144
Epoch 678: err = 0.001780452010279987: Win count = 464 Win ratio = 68.33578792341679
Epoch 679: err = 0.0013761165864707436: Win count = 465 Win ratio = 68.38235294117648
Epoch 680: err = 0.0011798315317719243: Win count = 466 Win ratio = 68.4287812041116
Epoch 681: err = 0.0015888244888628833: Win count = 466 Win ratio = 68.32844574780059
Epoch 682: err = 0.0017825497197918594: Win count = 467 Win ratio = 68.37481698389458
Epoch 683: err = 0.0013910464040236548: Win count = 468 Win ratio = 68.42105263157895
Epoch 684: err = 0.000926141336094588: Win count = 469 Win ratio = 68.46715328467154
Epoch 685: err = 0.0012039901048410684: Win count = 470 Win ratio = 68.5131195335277
Epoch 686: err = 0.00105904557130998: Win count = 471 Win ratio = 68.5589519650655
Epoch 687: err = 0.0008669111412018538: Win count = 472 Win ratio = 68.6046511627907
Epoch 688: err = 0.0011712055566022173: Win count = 473 Win ratio = 68.65021770682148
Epoch 689: err = 0.0007013580216153059: Win count = 474 Win ratio = 68.69565217391305
Epoch 690: err = 0.0008143602335621836: Win count = 475 Win ratio = 68.7409551374819
Epoch 691: err = 0.0007030424822005443: Win count = 476 Win ratio = 68.78612716763006
Epoch 692: err = 0.0008089173679763917: Win count = 477 Win ratio = 68.83116883116884
Epoch 693: err = 0.001161838792540948: Win count = 478 Win ratio = 68.87608069164266
Epoch 694: err = 0.0005747049362980761: Win count = 479 Win ratio = 68.92086330935253
Epoch 695: err = 0.0006537970475619659: Win count = 480 Win ratio = 68.96551724137932
Epoch 696: err = 0.0006782235941500403: Win count = 481 Win ratio = 69.01004304160688
Epoch 697: err = 0.0008597345222369768: Win count = 482 Win ratio = 69.05444126074498
Epoch 698: err = 0.0007955315450089984: Win count = 483 Win ratio = 69.09871244635193
Epoch 699: err = 0.00084668429553858: Win count = 484 Win ratio = 69.14285714285714
Epoch 700: err = 0.0006503501317638438: Win count = 484 Win ratio = 69.04422253922966
Epoch 701: err = 0.0008188547690224368: Win count = 485 Win ratio = 69.08831908831908
Epoch 702: err = 0.0007751389312034007: Win count = 486 Win ratio = 69.13229018492176
Epoch 703: err = 0.0005952297051408095: Win count = 487 Win ratio = 69.17613636363636
Epoch 704: err = 0.0005476305814227089: Win count = 488 Win ratio = 69.21985815602837
Epoch 705: err = 0.0005791563617094653: Win count = 489 Win ratio = 69.26345609065156
Epoch 706: err = 0.0005405196625360986: Win count = 490 Win ratio = 69.3069306930693
Epoch 707: err = 0.0006189280429680366: Win count = 491 Win ratio = 69.35028248587571
Epoch 708: err = 0.0006859029617771739: Win count = 492 Win ratio = 69.39351198871651
Epoch 709: err = 0.00047712979539937805: Win count = 493 Win ratio = 69.43661971830986
Epoch 710: err = 0.0007098739315551938: Win count = 494 Win ratio = 69.47960618846695
Epoch 711: err = 0.000790108821092872: Win count = 495 Win ratio = 69.52247191011236
Epoch 712: err = 0.0006964810145291267: Win count = 496 Win ratio = 69.56521739130434
Epoch 713: err = 0.0005682694099959917: Win count = 497 Win ratio = 69.6078431372549
Epoch 714: err = 0.0007282654078153428: Win count = 498 Win ratio = 69.65034965034965
Epoch 715: err = 0.0002983639296871843: Win count = 499 Win ratio = 69.6927374301676
Epoch 716: err = 0.0006568905482708942: Win count = 500 Win ratio = 69.7350069735007
Epoch 717: err = 0.0005383910975069739: Win count = 501 Win ratio = 69.77715877437326
Epoch 718: err = 0.0005606311606243253: Win count = 502 Win ratio = 69.81919332406119

Epoch 719: err = 0.0006541494585690089: Win count = 503 Win ratio = 69.86111111111111
Epoch 720: err = 0.0003669748257379979: Win count = 504 Win ratio = 69.90291262135922
Epoch 721: err = 0.0007196876722446177: Win count = 505 Win ratio = 69.94459833795014
Epoch 722: err = 0.000489078309328761: Win count = 506 Win ratio = 69.98616874135546
Epoch 723: err = 0.000531267451151507: Win count = 507 Win ratio = 70.02762430939227
Epoch 724: err = 0.0005805508517369162: Win count = 508 Win ratio = 70.06896551724138
Epoch 725: err = 0.0008155864925356582: Win count = 509 Win ratio = 70.11019283746556
Epoch 726: err = 0.0005531915594474413: Win count = 509 Win ratio = 70.01375515818432
Epoch 727: err = 0.00044749152402800974: Win count = 510 Win ratio = 70.05494505494505
Epoch 728: err = 0.0006438882301154081: Win count = 511 Win ratio = 70.0960219478738
Epoch 729: err = 0.0005969707526674028: Win count = 511 Win ratio = 70.0
Epoch 730: err = 0.0006401229402399622: Win count = 511 Win ratio = 69.90424076607387
Epoch 731: err = 0.0005291356173984241: Win count = 511 Win ratio = 69.80874316939891
Epoch 732: err = 0.00044995982716500293: Win count = 512 Win ratio = 69.84993178717599
Epoch 733: err = 0.00040073175841826014: Win count = 513 Win ratio = 69.89100817438693
Epoch 734: err = 0.00042448915792192565: Win count = 514 Win ratio = 69.93197278911565
Epoch 735: err = 0.0004737370982184075: Win count = 515 Win ratio = 69.97282608695652
Epoch 736: err = 0.0005506395027623512: Win count = 516 Win ratio = 70.01356852103122
Epoch 737: err = 0.0005590419496002141: Win count = 517 Win ratio = 70.05420054200542
Epoch 738: err = 0.0006176847091410309: Win count = 518 Win ratio = 70.09472259810555
Epoch 739: err = 0.0003614028282754589: Win count = 519 Win ratio = 70.13513513513514
Epoch 740: err = 0.0003738183186214883: Win count = 520 Win ratio = 70.17543859649122
Epoch 741: err = 0.00035730961826629937: Win count = 521 Win ratio = 70.21563342318059
Epoch 742: err = 0.0003165775906381896: Win count = 522 Win ratio = 70.2557200538358
Epoch 743: err = 0.0004249306202837033: Win count = 523 Win ratio = 70.29569892473118
Epoch 744: err = 0.00038331959331117105: Win count = 524 Win ratio = 70.33557046979865
Epoch 745: err = 0.0003143376961816102: Win count = 525 Win ratio = 70.37533512064343
Epoch 746: err = 0.00020008735555165913: Win count = 526 Win ratio = 70.41499330655957
Epoch 747: err = 0.0003389543962839525: Win count = 527 Win ratio = 70.45454545454545
Epoch 748: err = 0.00026166357019974384: Win count = 528 Win ratio = 70.4939919893191
Epoch 749: err = 0.00025057651873794384: Win count = 529 Win ratio = 70.53333333333333
Epoch 750: err = 0.0003308083669253392: Win count = 530 Win ratio = 70.57256990679095
Epoch 751: err = 0.00027369964755052933: Win count = 531 Win ratio = 70.61170212765957
Epoch 752: err = 0.0002450681658956455: Win count = 532 Win ratio = 70.6507304116866
Epoch 753: err = 0.00024367732476093806: Win count = 533 Win ratio = 70.6896551724138
Epoch 754: err = 0.00024965743978100363: Win count = 533 Win ratio = 70.59602649006622
Epoch 755: err = 0.0002272468809678685: Win count = 533 Win ratio = 70.5026455026455
Epoch 756: err = 0.0004078151159774279: Win count = 533 Win ratio = 70.40951122853369
Epoch 757: err = 0.0006179179890750675: Win count = 534 Win ratio = 70.44854881266491
Epoch 758: err = 0.0009886630850814981: Win count = 535 Win ratio = 70.4874835309618
Epoch 759: err = 0.0007333140001719585: Win count = 536 Win ratio = 70.52631578947368
Epoch 760: err = 0.0013320866819412913: Win count = 537 Win ratio = 70.56504599211564
Epoch 761: err = 0.0010840338036359753: Win count = 538 Win ratio = 70.60367454068242
Epoch 762: err = 0.0013763801689492539: Win count = 538 Win ratio = 70.51114023591087
Epoch 763: err = 0.0015370585097116418: Win count = 539 Win ratio = 70.54973821989529
Epoch 764: err = 0.0008350680400326382: Win count = 540 Win ratio = 70.58823529411765
Epoch 765: err = 0.0004840655565203633: Win count = 541 Win ratio = 70.62663185378591
Epoch 766: err = 0.000941119320486905: Win count = 542 Win ratio = 70.66492829204694
Epoch 767: err = 0.0006633296798099764: Win count = 543 Win ratio = 70.703125
Epoch 768: err = 0.0006226348523341585: Win count = 543 Win ratio = 70.6111833550065
Epoch 769: err = 0.000859038607814: Win count = 544 Win ratio = 70.64935064935065
Epoch 770: err = 0.0005979244797345018: Win count = 545 Win ratio = 70.68741893644618
Epoch 771: err = 0.0006933164258953184: Win count = 546 Win ratio = 70.72538860103627
Epoch 772: err = 0.0006304851922322996: Win count = 547 Win ratio = 70.76326002587322
Epoch 773: err = 0.0009080271229322534: Win count = 547 Win ratio = 70.67183462532299
Epoch 774: err = 0.0007667790487175807: Win count = 548 Win ratio = 70.70967741935485
Epoch 775: err = 0.00045950883395562414: Win count = 548 Win ratio = 70.61855670103093
Epoch 776: err = 0.0003285916209279094: Win count = 549 Win ratio = 70.65637065637065
Epoch 777: err = 0.000577794894525141: Win count = 549 Win ratio = 70.56555269922879
Epoch 778: err = 0.0006239455215109047: Win count = 550 Win ratio = 70.60333761232349

Epoch 779: err = 0.0006627390885114437: Win count = 551 Win ratio = 70.64102564102565
Epoch 780: err = 0.00043000456753361505: Win count = 552 Win ratio = 70.6786171574904
Epoch 781: err = 0.00046659887811983936: Win count = 553 Win ratio = 70.71611253196932
Epoch 782: err = 0.0004940240814903518: Win count = 554 Win ratio = 70.75351213282248
Epoch 783: err = 0.0006275452833506279: Win count = 555 Win ratio = 70.79081632653062
Epoch 784: err = 0.0004489301772991894: Win count = 556 Win ratio = 70.828025477707
Epoch 785: err = 0.0005073702450317796: Win count = 557 Win ratio = 70.86513994910942
Epoch 786: err = 0.0005147164138179505: Win count = 558 Win ratio = 70.90216010165184
Epoch 787: err = 0.00040288067611982115: Win count = 559 Win ratio = 70.93908629441624
Epoch 788: err = 0.0003216295408492442: Win count = 560 Win ratio = 70.97591888466414
Epoch 789: err = 0.0003197202768205898: Win count = 561 Win ratio = 71.0126582278481
Epoch 790: err = 0.00027915268037759233: Win count = 562 Win ratio = 71.04930467762325
Epoch 791: err = 0.00046817903239571024: Win count = 563 Win ratio = 71.08585858585859
Epoch 792: err = 0.00034705427970038727: Win count = 564 Win ratio = 71.12232030264816
Epoch 793: err = 0.00037633102510881145: Win count = 565 Win ratio = 71.15869017632241
Epoch 794: err = 0.0003329607188788941: Win count = 566 Win ratio = 71.19496855345912
Epoch 795: err = 0.00025410450052731903: Win count = 567 Win ratio = 71.23115577889448
Epoch 796: err = 0.00040873733632906806: Win count = 568 Win ratio = 71.267252195734
Epoch 797: err = 0.0003541835340001853: Win count = 569 Win ratio = 71.30325814536342
Epoch 798: err = 0.0002780111644824501: Win count = 570 Win ratio = 71.33917396745932
Epoch 799: err = 0.00032884501342778094: Win count = 571 Win ratio = 71.375
Epoch 800: err = 0.00026417588014737703: Win count = 572 Win ratio = 71.4107365792759
Epoch 801: err = 0.0002290705633640755: Win count = 573 Win ratio = 71.44638403990025
Epoch 802: err = 0.00029922815883764997: Win count = 574 Win ratio = 71.48194271481943
Epoch 803: err = 0.0002739342180575477: Win count = 575 Win ratio = 71.51741293532339
Epoch 804: err = 0.00023285312545340275: Win count = 576 Win ratio = 71.55279503105591
Epoch 805: err = 0.00020906952886434738: Win count = 577 Win ratio = 71.58808933002481
Epoch 806: err = 0.0005427713040262461: Win count = 578 Win ratio = 71.62329615861215
Epoch 807: err = 0.0004104418685528799: Win count = 579 Win ratio = 71.65841584158416
Epoch 808: err = 0.00039468796603614464: Win count = 580 Win ratio = 71.69344870210136
Epoch 809: err = 0.0005273580372886499: Win count = 581 Win ratio = 71.7283950617284
Epoch 810: err = 0.00024346262671315344: Win count = 582 Win ratio = 71.7632552404439
Epoch 811: err = 0.0002500019036233425: Win count = 583 Win ratio = 71.79802955665025
Epoch 812: err = 0.00020192346710246056: Win count = 584 Win ratio = 71.83271832718327
Epoch 813: err = 0.00035100768127449555: Win count = 585 Win ratio = 71.86732186732186
Epoch 814: err = 0.00027512638007465284: Win count = 586 Win ratio = 71.90184049079754
Epoch 815: err = 0.0002496250417607371: Win count = 587 Win ratio = 71.93627450980392
Epoch 816: err = 0.0006654170283582062: Win count = 588 Win ratio = 71.97062423500611
Epoch 817: err = 0.0005192577464185888: Win count = 589 Win ratio = 72.00488997555013
Epoch 818: err = 0.000384766519346158: Win count = 589 Win ratio = 71.91697191697192
Epoch 819: err = 0.0003835280349449022: Win count = 590 Win ratio = 71.95121951219512
Epoch 820: err = 0.00030489818527712487: Win count = 591 Win ratio = 71.98538367844093
Epoch 821: err = 0.0005812072067783447: Win count = 592 Win ratio = 72.01946472019465
Epoch 822: err = 0.00047055992763489485: Win count = 593 Win ratio = 72.05346294046173
Epoch 823: err = 0.00044380083272699267: Win count = 594 Win ratio = 72.0873786407767
Epoch 824: err = 0.0004280021876184037: Win count = 595 Win ratio = 72.12121212121212
Epoch 825: err = 0.0003340921593917301: Win count = 596 Win ratio = 72.15496368038741
Epoch 826: err = 0.00036285144869907526: Win count = 597 Win ratio = 72.18863361547763
Epoch 827: err = 0.00017994347217609175: Win count = 598 Win ratio = 72.22222222222221
Epoch 828: err = 0.00021781687701150076: Win count = 599 Win ratio = 72.25572979493366
Epoch 829: err = 0.00036239811925042886: Win count = 600 Win ratio = 72.28915662650603
Epoch 830: err = 0.00016829873675305862: Win count = 600 Win ratio = 72.20216606498195
Epoch 831: err = 0.0002125942992279306: Win count = 601 Win ratio = 72.23557692307693
Epoch 832: err = 0.00016826233422762016: Win count = 602 Win ratio = 72.26890756302521
Epoch 833: err = 0.00015399709627672564: Win count = 602 Win ratio = 72.18225419664267
Epoch 834: err = 0.00018182492931373417: Win count = 603 Win ratio = 72.21556886227545
Epoch 835: err = 0.00017742711543178302: Win count = 604 Win ratio = 72.24880382775119
Epoch 836: err = 0.000148977564094821: Win count = 605 Win ratio = 72.28195937873357
Epoch 837: err = 0.0001598364156052412: Win count = 606 Win ratio = 72.31503579952268
Epoch 838: err = 0.00020774149379576556: Win count = 607 Win ratio = 72.34803337306317

Epoch 839: err = 0.000174569280261494: Win count = 608 Win ratio = 72.38095238095238
Epoch 840: err = 0.00021195754015934654: Win count = 609 Win ratio = 72.41379310344827
Epoch 841: err = 0.0003143737594655249: Win count = 610 Win ratio = 72.44655581947742
Epoch 842: err = 0.00016186589027711307: Win count = 611 Win ratio = 72.47924080664293
Epoch 843: err = 0.00012622853319044225: Win count = 612 Win ratio = 72.51184834123224
Epoch 844: err = 0.0003185139976267237: Win count = 613 Win ratio = 72.54437869822485
Epoch 845: err = 0.000333482731548429: Win count = 614 Win ratio = 72.57683215130024
Epoch 846: err = 0.0004019930202048272: Win count = 615 Win ratio = 72.60920897284534
Epoch 847: err = 0.00017852653672889573: Win count = 616 Win ratio = 72.64150943396226
Epoch 848: err = 0.00023875311671872623: Win count = 617 Win ratio = 72.67373380447584
Epoch 849: err = 0.00027189781849301653: Win count = 618 Win ratio = 72.70588235294117
Epoch 850: err = 0.00016194933687074808: Win count = 619 Win ratio = 72.73795534665099
Epoch 851: err = 0.00022904109664523276: Win count = 620 Win ratio = 72.76995305164318
Epoch 852: err = 0.0002174035844291211: Win count = 621 Win ratio = 72.80187573270808
Epoch 853: err = 0.00022091829578130273: Win count = 622 Win ratio = 72.8337236533958
Epoch 854: err = 0.00020091994792892365: Win count = 623 Win ratio = 72.86549707602339
Epoch 855: err = 0.00019030733028557734: Win count = 623 Win ratio = 72.7803738317757
Epoch 856: err = 0.0001803174645829131: Win count = 624 Win ratio = 72.81213535589265
Epoch 857: err = 0.0001562539227961679: Win count = 625 Win ratio = 72.84382284382285
Epoch 858: err = 0.0001594230288901599: Win count = 626 Win ratio = 72.8754365541327
Epoch 859: err = 0.00016744532604207052: Win count = 627 Win ratio = 72.90697674418605
Epoch 860: err = 0.0001716822571324883: Win count = 628 Win ratio = 72.93844367015099
Epoch 861: err = 0.00014406250875254045: Win count = 629 Win ratio = 72.96983758700696
Epoch 862: err = 0.0001234585884049011: Win count = 630 Win ratio = 73.00115874855156
Epoch 863: err = 0.0001359048660560802: Win count = 631 Win ratio = 73.0324074074074
Epoch 864: err = 0.00023919795603433158: Win count = 632 Win ratio = 73.06358381502889
Epoch 865: err = 0.00018628767429618165: Win count = 633 Win ratio = 73.094688221709
Epoch 866: err = 0.00016689623680576915: Win count = 634 Win ratio = 73.12572087658593
Epoch 867: err = 0.00021640106660925085: Win count = 635 Win ratio = 73.15668202764977
Epoch 868: err = 0.00010291921716998331: Win count = 636 Win ratio = 73.18757192174914
Epoch 869: err = 0.0001613021031516837: Win count = 637 Win ratio = 73.2183908045977
Epoch 870: err = 0.00017750901997715118: Win count = 638 Win ratio = 73.24913892078071
Epoch 871: err = 0.00012368236548354616: Win count = 639 Win ratio = 73.27981651376146
Epoch 872: err = 0.00016661978770571295: Win count = 640 Win ratio = 73.31042382588774
Epoch 873: err = 0.0001626443372515496: Win count = 641 Win ratio = 73.34096109839817
Epoch 874: err = 0.0001752430939632177: Win count = 642 Win ratio = 73.37142857142858
Epoch 875: err = 0.00016956915033006226: Win count = 643 Win ratio = 73.40182648401826
Epoch 876: err = 0.00013557683905673912: Win count = 644 Win ratio = 73.4321550741163
Epoch 877: err = 8.757458545005647e-05: Win count = 645 Win ratio = 73.4624145785877
Epoch 878: err = 0.00010057286272058263: Win count = 646 Win ratio = 73.49260523321956
Epoch 879: err = 0.00043788478569695144: Win count = 647 Win ratio = 73.52272727272727
Epoch 880: err = 0.0003243773999201949: Win count = 648 Win ratio = 73.55278093076049
Epoch 881: err = 0.0002166425583709497: Win count = 649 Win ratio = 73.58276643990929
Epoch 882: err = 0.00023473949386243476: Win count = 650 Win ratio = 73.61268403171007
Epoch 883: err = 0.00025104122414632: Win count = 651 Win ratio = 73.64253393665159
Epoch 884: err = 0.0002924520977103384: Win count = 652 Win ratio = 73.67231638418079
Epoch 885: err = 0.0001221091924890061: Win count = 653 Win ratio = 73.7020316027088
Epoch 886: err = 0.00020391595717228483: Win count = 654 Win ratio = 73.73167981961669
Epoch 887: err = 0.00020609735383914085: Win count = 655 Win ratio = 73.76126126126125
Epoch 888: err = 0.00023689644422120182: Win count = 655 Win ratio = 73.67829021372329
Epoch 889: err = 0.00013796327857562574: Win count = 656 Win ratio = 73.70786516853933
Epoch 890: err = 9.172611999019864e-05: Win count = 656 Win ratio = 73.62514029180696
Epoch 891: err = 0.0001798126618268725: Win count = 657 Win ratio = 73.65470852017937
Epoch 892: err = 0.000208109435334336: Win count = 658 Win ratio = 73.68421052631578
Epoch 893: err = 0.0001848524589149747: Win count = 659 Win ratio = 73.71364653243847
Epoch 894: err = 0.0002691387253435096: Win count = 660 Win ratio = 73.74301675977654
Epoch 895: err = 0.00012087272170901997: Win count = 661 Win ratio = 73.77232142857143
Epoch 896: err = 0.0001127467157857609: Win count = 662 Win ratio = 73.8015607580825
Epoch 897: err = 0.00010284115387548809: Win count = 663 Win ratio = 73.83073496659243
Epoch 898: err = 8.241075374826323e-05: Win count = 663 Win ratio = 73.74860956618465

Epoch 899: err = 9.541422650727327e-05: Win count = 664 Win ratio = 73.77777777777777
Epoch 900: err = 8.271390834124759e-05: Win count = 665 Win ratio = 73.80688124306326
Epoch 901: err = 5.2229029051886755e-05: Win count = 666 Win ratio = 73.8359201773836
Epoch 902: err = 7.177584893724998e-05: Win count = 667 Win ratio = 73.86489479512736
Epoch 903: err = 0.00010185643077420536: Win count = 668 Win ratio = 73.89380530973452
Epoch 904: err = 0.00012351344003036502: Win count = 669 Win ratio = 73.92265193370166
Epoch 905: err = 0.00014954235211916966: Win count = 670 Win ratio = 73.9514348785872
Epoch 906: err = 0.00011261659074079944: Win count = 671 Win ratio = 73.98015435501654
Epoch 907: err = 7.740453975202399e-05: Win count = 672 Win ratio = 74.00881057268722
Epoch 908: err = 0.0001161538662017847: Win count = 673 Win ratio = 74.03740374037405
Epoch 909: err = 0.00012518855919552152: Win count = 674 Win ratio = 74.06593406593407
Epoch 910: err = 0.00011027943219232839: Win count = 675 Win ratio = 74.09440175631174
Epoch 911: err = 6.495894695035531e-05: Win count = 675 Win ratio = 74.01315789473685
Epoch 912: err = 7.827733975318552e-05: Win count = 676 Win ratio = 74.04162102957284
Epoch 913: err = 0.00011982736032223329: Win count = 677 Win ratio = 74.07002188183807
Epoch 914: err = 6.964442218304612e-05: Win count = 678 Win ratio = 74.09836065573771
Epoch 915: err = 7.537601004514727e-05: Win count = 679 Win ratio = 74.12663755458514
Epoch 916: err = 9.468638609178015e-05: Win count = 680 Win ratio = 74.15485278080698
Epoch 917: err = 0.00010027008920587832: Win count = 681 Win ratio = 74.18300653594771
Epoch 918: err = 0.00010822170361279859: Win count = 682 Win ratio = 74.21109902067464
Epoch 919: err = 9.698122994450387e-05: Win count = 683 Win ratio = 74.23913043478261
Epoch 920: err = 0.0001090472360374406: Win count = 684 Win ratio = 74.2671009771987
Epoch 921: err = 0.00011276533132331679: Win count = 685 Win ratio = 74.29501084598698
Epoch 922: err = 7.916074127933825e-05: Win count = 686 Win ratio = 74.3228602383532
Epoch 923: err = 0.00012080180567863863: Win count = 687 Win ratio = 74.35064935064936
Epoch 924: err = 6.85718246131728e-05: Win count = 688 Win ratio = 74.37837837837837
Epoch 925: err = 7.251457464008126e-05: Win count = 688 Win ratio = 74.29805615550757
Epoch 926: err = 7.622484690728015e-05: Win count = 689 Win ratio = 74.32578209277239
Epoch 927: err = 5.653410426020855e-05: Win count = 690 Win ratio = 74.35344827586206
Epoch 928: err = 7.48114293855906e-05: Win count = 691 Win ratio = 74.3810548977395
Epoch 929: err = 7.14912803232437e-05: Win count = 692 Win ratio = 74.40860215053763
Epoch 930: err = 8.42550907691475e-05: Win count = 693 Win ratio = 74.43609022556392
Epoch 931: err = 8.462674259135383e-05: Win count = 694 Win ratio = 74.46351931330472
Epoch 932: err = 6.757744131391519e-05: Win count = 695 Win ratio = 74.49088960342979
Epoch 933: err = 5.1170476126571884e-05: Win count = 695 Win ratio = 74.41113490364025
Epoch 934: err = 4.9135634526464855e-05: Win count = 696 Win ratio = 74.43850267379679
Epoch 935: err = 5.3704062793258345e-05: Win count = 697 Win ratio = 74.46581196581197
Epoch 936: err = 3.931281617042259e-05: Win count = 698 Win ratio = 74.49306296691569
Epoch 937: err = 4.3457855099404696e-05: Win count = 699 Win ratio = 74.52025586353945
Epoch 938: err = 4.9794021379057085e-05: Win count = 700 Win ratio = 74.54739084132055
Epoch 939: err = 6.721297086187406e-05: Win count = 701 Win ratio = 74.57446808510639
Epoch 940: err = 9.163551385427127e-05: Win count = 702 Win ratio = 74.60148777895856
Epoch 941: err = 6.830931920376315e-05: Win count = 702 Win ratio = 74.52229299363057
Epoch 942: err = 6.325105914584128e-05: Win count = 703 Win ratio = 74.5493107104984
Epoch 943: err = 8.156054536812007e-05: Win count = 703 Win ratio = 74.47033898305084
Epoch 944: err = 5.28025097992213e-05: Win count = 704 Win ratio = 74.4973544973545
Epoch 945: err = 0.00012513830688476446: Win count = 705 Win ratio = 74.52431289640592
Epoch 946: err = 0.0001545947275189974: Win count = 706 Win ratio = 74.55121436114044
Epoch 947: err = 0.000163027951884942: Win count = 707 Win ratio = 74.57805907172997
Epoch 948: err = 0.00026781218093674397: Win count = 708 Win ratio = 74.60484720758693
Epoch 949: err = 0.00010423738012832473: Win count = 709 Win ratio = 74.63157894736842
Epoch 950: err = 0.00018421944105284638: Win count = 710 Win ratio = 74.65825446898002
Epoch 951: err = 6.89025134761323e-05: Win count = 710 Win ratio = 74.57983193277312
Epoch 952: err = 0.000108004959884056: Win count = 711 Win ratio = 74.60650577124869
Epoch 953: err = 0.00018618544072523946: Win count = 712 Win ratio = 74.63312368972747
Epoch 954: err = 0.00027319765376887517: Win count = 713 Win ratio = 74.65968586387434
Epoch 955: err = 0.00013729918146054843: Win count = 714 Win ratio = 74.68619246861925
Epoch 956: err = 8.804263734418782e-05: Win count = 715 Win ratio = 74.71264367816092
Epoch 957: err = 7.678385873077787e-05: Win count = 716 Win ratio = 74.73903966597078
Epoch 958: err = 0.0002266653127662721: Win count = 717 Win ratio = 74.76538060479666

```
Epoch 959: err = 0.00019925631022488233: Win count = 718 Win ratio = 74.79166666666667
Epoch 960: err = 0.00019926821323679178: Win count = 719 Win ratio = 74.81789802289282
Epoch 961: err = 0.00010260959015795379: Win count = 720 Win ratio = 74.84407484407485
Epoch 962: err = 0.00023711667745374143: Win count = 721 Win ratio = 74.87019730010385
Epoch 963: err = 0.00011897974582097959: Win count = 722 Win ratio = 74.89626556016597
Epoch 964: err = 0.00013801545992464526: Win count = 723 Win ratio = 74.92227979274611
Epoch 965: err = 0.00013444297474052291: Win count = 724 Win ratio = 74.94824016563147
Epoch 966: err = 9.964644414139912e-05: Win count = 725 Win ratio = 74.9741468459152
Epoch 967: err = 0.00021224367901595542: Win count = 726 Win ratio = 75.0
Epoch 968: err = 7.08066445440636e-05: Win count = 727 Win ratio = 75.02579979360165
Epoch 969: err = 0.00019009905190614518: Win count = 728 Win ratio = 75.05154639175258
Epoch 970: err = 8.874555896909442e-05: Win count = 729 Win ratio = 75.07723995880535
Epoch 971: err = 0.00016150899273270625: Win count = 730 Win ratio = 75.1028806584362
Epoch 972: err = 0.00016907053804970928: Win count = 731 Win ratio = 75.12846865364851
Epoch 973: err = 0.00017791748814488528: Win count = 732 Win ratio = 75.15400410677618
Epoch 974: err = 5.953756976850855e-05: Win count = 733 Win ratio = 75.17948717948718
Epoch 975: err = 0.00017658745218795957: Win count = 734 Win ratio = 75.20491803278688
Epoch 976: err = 0.00014619466310250573: Win count = 735 Win ratio = 75.23029682702149
Epoch 977: err = 0.00012431159984771512: Win count = 736 Win ratio = 75.25562372188139
Epoch 978: err = 0.00011358083042978251: Win count = 737 Win ratio = 75.28089887640449
Epoch 979: err = 0.0001223486060553114: Win count = 738 Win ratio = 75.3061224489796
Epoch 980: err = 0.00011065321177738952: Win count = 739 Win ratio = 75.33129459734964
Epoch 981: err = 9.667016047387733e-05: Win count = 740 Win ratio = 75.35641547861506
Epoch 982: err = 0.00041453260428170324: Win count = 741 Win ratio = 75.38148524923703
Epoch 983: err = 0.00014785600797040388: Win count = 742 Win ratio = 75.40650406504065
Epoch 984: err = 0.00010620393527460692: Win count = 743 Win ratio = 75.43147208121827
Epoch 985: err = 0.0003430983842918067: Win count = 744 Win ratio = 75.45638945233266
Epoch 986: err = 0.00016104550604723045: Win count = 745 Win ratio = 75.48125633232016
Epoch 987: err = 9.831252100411803e-05: Win count = 745 Win ratio = 75.40485829959515
Epoch 988: err = 0.00029768000922558713: Win count = 746 Win ratio = 75.42972699696662
Epoch 989: err = 0.0002987210236824467: Win count = 747 Win ratio = 75.45454545454545
Epoch 990: err = 0.0004201586261842749: Win count = 748 Win ratio = 75.47931382441978
Epoch 991: err = 0.0002527479459786264: Win count = 749 Win ratio = 75.50403225806451
Epoch 992: err = 0.00024656833375047427: Win count = 750 Win ratio = 75.52870090634441
Epoch 993: err = 0.00013458837838697946: Win count = 751 Win ratio = 75.55331991951711
Epoch 994: err = 8.707413871889003e-05: Win count = 752 Win ratio = 75.57788944723617
Epoch 995: err = 0.0003607154412748059: Win count = 753 Win ratio = 75.6024096385542
Epoch 996: err = 6.615783513552742e-05: Win count = 754 Win ratio = 75.62688064192578
Epoch 997: err = 0.00011952717568419757: Win count = 755 Win ratio = 75.65130260521042
Epoch 998: err = 0.0002599767904030159: Win count = 756 Win ratio = 75.67567567567568
Epoch 999: err = 0.0001271523560717469: Win count = 757 Win ratio = 75.7
Epoch 1000: err = 0.0002586702630651416: Win count = 758 Win ratio = 75.72427572427573
Finish
```

An exception has occurred, use %tb to see the full traceback.

SystemExit

플레이 & 시각화

import

```
In [50]: from IPython import display
import matplotlib.pyplot as plt
import matplotlib.patches as patches
import pylab as pl
```

```
import time
import tensorflow as tf
import math
import os
```

파라미터 설정

```
In [51]: gridSize = 10 # 게임을 플레이할 격자 크기
maxGames = 10 # 게임을 플레이할 횟수
env = CatchEnvironment(gridSize) # 환경 설정
winCount = 0 # 게임 승리 횟수
loseCount = 0 # 게임 패배 횟수
numberOfGames = 0 # 게임 실행 횟수
```

화면 설정

ground = 1 plot = pl.figure(figsize=(12,12))

크기가 12 x 12인 figure를 생성

axis = plot.add_subplot(111, aspect='equal')

만들어진 figure에 subplot을 추가해준다

111은 1행 1열의 1번이라는 의미 (figure 전체) aspect='equal'은 축의 비율을 갖게 만들어준다

axis.set_xlim([-1, 12])

x축의 범위를 -1 ~ 12까지로 설정

axis.set_ylim([0, 12])

y축의 범위를 0 ~ 12까지로 설정

drawState 함수

drawState 함수는 현재 상태를 그려준다

모델을 사용하여 게임 플레이

saver = tf.compat.v1.train.Saver() with tf.compat.v1.Session() as sess: 세션을 실행시켜준다

세션을 나타내는 객체는 sess가 된다

saver.restore(sess, 'model')

저장했었던 model을 불러온다

print('saved model is loaded!')

while (numberOfGames < maxGames):

만약 numberOfGames가 maxGames 보다 작다면 게임을 플레이한다

numberOfGames = numberOfGames + 1

numberOfGames의 값을 1 더해준다

isGameOver = False

fruitRow, fruitColumn, basket = env.reset()

currentState = env.observe()

drawState(fruitRow, fruitColumn, basket)

메인함수에서와 마찬가지로 값들을 초기화 시켜준다

여기서 다른점은 reset 함수를 사용해 리셋한 값을 불러와 drawState 함수에 넣어 게임화면을 그려준다

while (isGameOver != True):

만약 게임이 끝나지 않았다면 (만약 게임이 끝났다면 while문이 풀리면서 위쪽 while문으로 이동해 게임을 재시작하게 된다)

q = sess.run(output_layer, feed_dict={X: currentState})

위에서 정의했던 딥러닝에 currentState를 입력값으로 넣어 돌려 q에 넣어준다

index = q.argmax()

action = index + 1

딥러닝을 돌려 얻은 q값들 중 가장 큰 값에 1을 더하여 action으로 설정해준다

nextState, reward, gameOver, stateInfo = env.act(action)

act 함수를 사용해 nextState, reward, gameOver, stateInfo를 얻어낸다

여기까지는 메인함수와 똑같이 작동한다

fruitRow = stateInfo[0]

fruitColumn = stateInfo[1]

basket = stateInfo[2]

여기서는 메인함수와 다르게 화면에 표시할 정보가 필요하기 때문에 stateInfo에서 각 정보들을 가져온다

fruitRow(블록의 행번호)

fruitColumn(블록의 열번호)

basket(바의 위치)

if (reward == 1):

winCount = winCount + 1

elif (reward == -1):

loseCount = loseCount + 1

만약 reward가 1이라면 winCount를 1 더해주고 reward가 -1이라면 loseCount를 1 더해준다

currentState = nextState

현재 상태를 다음 상태(act함수를 실행한)로 업데이트 해준다

isGameOver = gameOver

게임 종료 여부를 업데이트 시켜준다

drawState(fruitRow, fruitColumn, basket)

아까 가져온 정보들로 drawState 함수를 사용해 게임 화면을 그려준다

time.sleep(0.4)

다음 행동을 하기 전에 딜레이를 준다

display.clear_output(wait=True)

마지막 결과를 새로운 화면으로 표시하는 것이 아닌 그대로 표시하기 위해 사용

```
In [52]: # 화면 설정
ground = 1
plot = pl.figure(figsize=(12,12))
axis = plot.add_subplot(111, aspect='equal')
axis.set_xlim([-1, 12])
axis.set_ylim([0, 12])

# drawState 함수
def drawState(fruitRow, fruitColumn, basket):
    global gridSize
    fruitX = fruitColumn
    fruitY = (gridSize - fruitRow + 1)
    statusTitle = "Wins: " + str(winCount) + " Losses: " + str(loseCount) + " TotalGame: " + str(numberOfGames)
    axis.set_title(statusTitle, fontsize=30)
    for p in [
        patches.Rectangle(
            ((ground - 1), (ground)), 11, 10, facecolor="#000000"      # Black
        ),
        patches.Rectangle(
            (basket - 1, ground), 2, 0.5, facecolor="#FF0000"      # No background
        ),
        patches.Rectangle(
            (fruitX - 0.5, fruitY - 0.5), 1, 1, facecolor="#FF0000"      # red
        ),
    ]:
        axis.add_patch(p)
    display.clear_output(wait=True)
    display.display(pl.gcf())

# 모델을 사용하여 게임 플레이
saver = tf.compat.v1.train.Saver()
with tf.compat.v1.Session() as sess:
    # model을 불러온다
    saver.restore(sess, 'model')
    print('saved model is loaded!')

    while (numberOfGames < maxGames):
        numberOfGames = numberOfGames + 1

        # 초기화
        isGameOver = False
        fruitRow, fruitColumn, basket = env.reset()
        currentState = env.observe()
```

```
drawState(fruitRow, fruitColumn, basket)
```

```
while (isGameOver != True):
```

```
    q = sess.run(output_layer, feed_dict={X: currentState})
```

```
    # 가장 큰 Q값을 사용
```

```
    index = q.argmax()
```

```
    action = index + 1
```

```
    nextState, reward, gameOver, stateInfo = env.act(action)
```

```
    fruitRow = stateInfo[0]
```

```
    fruitColumn = stateInfo[1]
```

```
    basket = stateInfo[2]
```

```
    # 게임 결과를 카운트
```

```
    if (reward == 1):
```

```
        winCount = winCount + 1
```

```
    elif (reward == -1):
```

```
        loseCount = loseCount + 1
```

```
    currentState = nextState
```

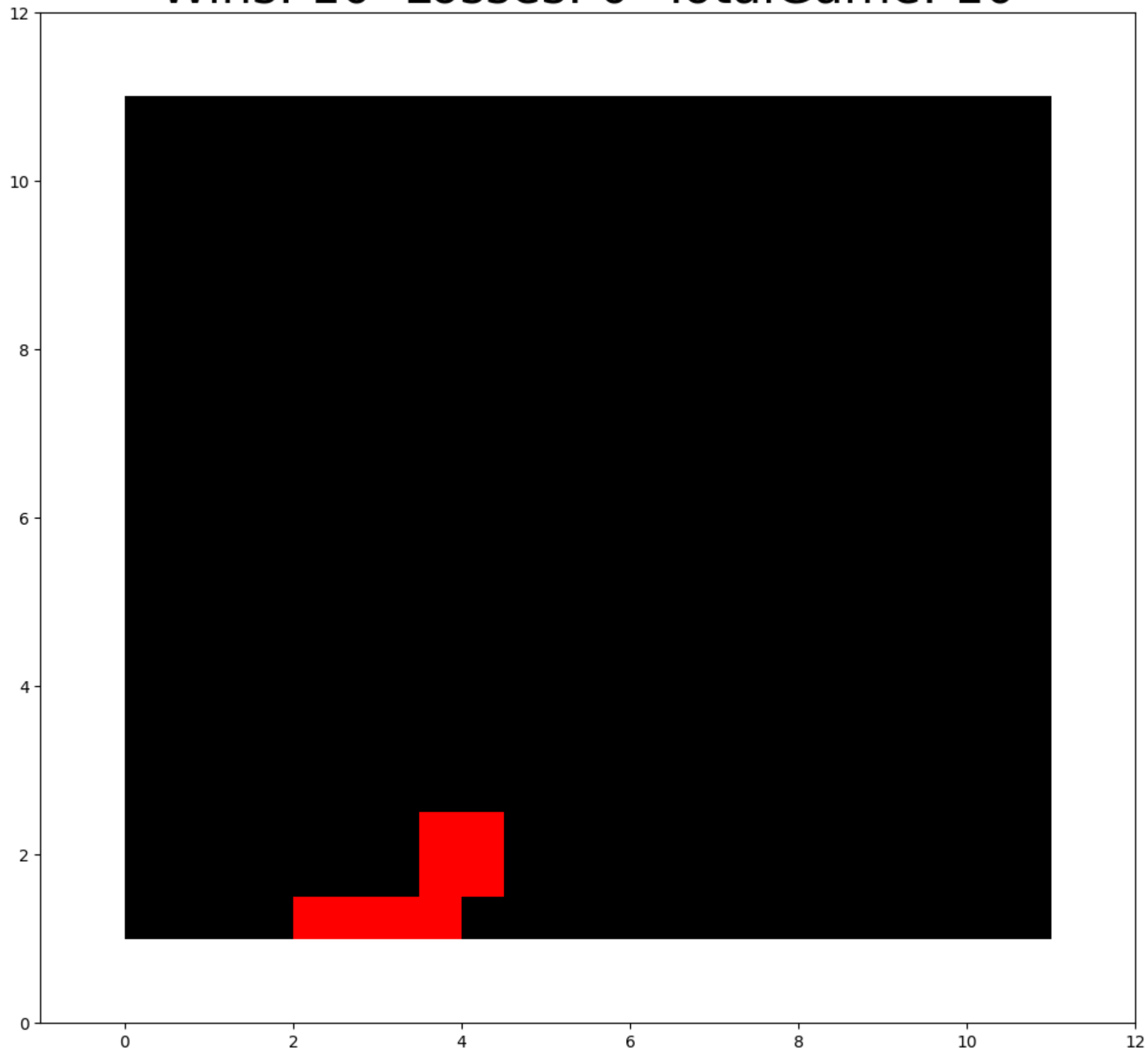
```
    isGameOver = gameOver
```

```
    drawState(fruitRow, fruitColumn, basket)
```

```
    time.sleep(0.4)
```

```
display.clear_output(wait=True)
```

Wins: 10 Losses: 0 TotalGame: 10



정리

여기서는 DQN을 이용하여 간단한 블럭 받기 게임을 학습시켰습니다

우선 epsilon을 사용하여 랜덤으로 행동할지 Q값에 따라 행동할지 결정합니다

랜덤으로 행동을 할 때에는 action값을 랜덤으로 주고 Q값에 따라 행동을 할 때에는 신경망에 현재 상태를 입력값으로 주고 나온 결과값들 중에서 가장 큰 값을 action값으로 설정합니다
랜덤으로 행동할 확률인 epsilon은 학습이 진행될 수록 감소합니다

그렇게 결정된 action을 수행하고 메모리에 action을 수행한 결과를 저장해 줍니다

그 다음 메모리에서 데이터 묶음인 batch를 가져와 학습을 시켜줍니다

inputs에는 입력값들이 들어있고 targets에는 보상값과 다음 상태의 최대 Q값이 더해진 값이 들어있습니다

이제 이 inputs값이 딥러닝을 거쳐 나온 결과값과 targets값과의 오차가 최소가 될 수 있도록 경사하강법을 이용하여 학습시켜줍니다

이러한 학습을 계속하게 되면 오차가 줄어들고 원하는 결과값을 얻을 수 있습니다