Computational Intelligence - Quarto

Author

Muhammad Sarib Khan s298885

Collaborators

To be completely candid, since there was no one for this session that I could team up with, I contacted my friend, Angelica Ferlin (an erasmus exchange student), who appeared in the January session and discussed my ideas with her.

Angelica Ferlin - Discussed possible solutions/problems and helped me with the management of Q tables

Resources

Links

- https://stackoverflow.com/questions/10016352/convert-numpy-array-to-tuple If there's an error with the Q-table because numpy arrays are not hashable, you can resolve it by converting the numpy array into a tuple.
- https://stackoverflow.com/questions/4901815/object-of-custom-type-as-dictionary-key I've found a solution for QTableKey here, where instead of using a tuple, an object is used as the key.
- https://realpython.com/python-is-identity-vs-equality/ I encountered a memory issue with the QTableKey, and this link clarified that I needed to implement an eq() method to enable comparison when hashing.
- Link to final repository: https://github.com/yurnero14/Quarto-Final
- Link to course repo: https://github.com/yurnero14/Computational-Intelligence-Muhammad-Sarib-Khan

Code Development

In the beginning of the project, Angelica did some research on different possible algorithms and Leonor studied the implementation of Minimax. However, after careful consideration, it was decided to implement Reinforcement Learning instead. I took this decision because I am much more confident implementing an RL based solution since I have done a few mini projects on my own in the past based on RL. Additionally, it was anticipated that Minimax would require a significant amount of time, given the large number of possible states in the game of Quarto.

In the process of developing the Reinforcement Learning strategy, a substantial amount of knowledge was gained from the concerned lectures. In all honesty, I am not the best in coding so Angelica helped giving me the direction to develop an algorithm to code for Q-learning.

Regarding the Q-table, the decision was made to combine the selected piece and the placement of the piece into a single move. Consequently, the key for the Q-table was structured as a tuple, incorporating both the current state (comprising the board array and the chosen piece) and the move itself. The Q-values were then associated with this key. To make the current state usable as part of the key, specific hash and equality (eq) functions had to be implemented. This approach was derived from a previously mentioned source.

In addition, following the creation of the RL agent with the goal of training it, another agent was intentionally designed to make poor decisions. This deliberate choice allowed the agent to undergo initial training with the "bad" agent before transitioning to random actions, thereby smoothing the overall learning process.

The ExtendedQuarto class was developed to enhance the integration of the custom-written code with the provided libraries. This class facilitated the ability to switch the current player, a critical aspect for the implemented logic and auxiliary functions to function smoothly.

Following the agent's training, the program stores the results in a file. This document is subsequently utilized by the trained agent class to read the Q-table and apply moves based on the knowledge acquired during training. Unfortunately, some setbacks were encountered due to the size of the file and using the pickle library. This library was generating a file that was too big to simply upload it on github. Hence, by using HFL (Large File Storage), a solution was found.

In addition, in order to get these results, the agent was trained with Genetic Algorith.

Code Map

- *extendedQuarto.py* An extended version of Quarto was created to incorporate additional functionality and features.
- *testQuarto.py* An extended version of Quarto was developed to enable more comprehensive move testing and evaluation.
- *rl.py* A Reinforcement Learning Agent and a corresponding Class designed to be used as a Key for the Q-table were implemented in the project.
- *opponent_agents.py* An intentionally designed agent was created to deliberately make suboptimal decisions as part of the training or testing process.
- train_q_learner.py includes a function that facilitates running a game between the Q-Learner and an opponent agent. Additionally, there's a strategy in place to guide the Q-Learner's learning process during these games.
- trained rl.py Class including the trained RL
- *q_table_1.pickle* The Q-table is saved in a file with the following parameters:
 - Number of games played: 1000
 - Games added per opponent: 500
 - Exploration rate decreases by 0.05 every 100th game.
- *q_table_2.pickle* The Q-Learner is saved in a file with the specified parameters:
 - Number of games played: 2000

- Games added per opponent: 700
- Exploration rate decreases by 0.05 every 200th game.
- $q_table_3.pickle$ The Q-Learner is saved in a file with the following parameters:
 - Number of games played: 3000
 - Games added per opponent: 700
 - Exploration rate decreases by 0.05 every 300th game.

How to run the code

- In train_q_learner.py file, number of games can be set at line 158 of the file (3rd parameter). At line 108, the number of games after which you want to decrease exploration rate can be change
- In rl.py, at line 273, name the pickle file you want to save the q-table in
- in trained_rl, at line 64, keep the name of the file to read same as the last bullet point
- run python main.py

Source code of project:

generate_agent.py:

```
import random
import quarto

class EvolvedAgent(quarto.Player):
    '''Evolved agent using the GA approach'''

def __init__(self, quarto: quarto.Quarto) -> None:
    super().__init__(quarto)
    self.board = quarto
    self.active_pieces = [i for i in range(16)]

    # Generate the probability of picking each rule for picking and

placing
    pick_prob_not_normalized = [random.random() for i in range(3)]
    pick_prob = [x/sum(pick_prob_not_normalized) for x in

pick_prob_not_normalized]
    self.pick_probability_1 = pick_prob[0]
    self.pick_probability_2 = pick_prob[1]
    self.pick_probability_3 = pick_prob[2]

    place_probability_not_normalized = [random.random() for i in

range(6)]
    place_probability_not_normalized = [random.random() for x in

place_probability_not_normalized]
    self.place_probability_1 = place_prob[0]
    self.place_probability_2 = place_prob[1]
    self.place_probability_3 = place_prob[2]
    self.place_probability_4 = place_prob[3]
    self.place_probability_6 = place_prob[4]
    self.place_probability_6 = place_prob[5]
```

```
def get place prob(self):
```

```
self.place probability 5:
```

```
def piece attribute dict(self) -> dict:
           dominating result = self.dominating(row, line counter)
self.dom line dict[str(dominating result[0])].append(dominating result)
           dominating result = self.dominating(board[:,i], line counter)
self.dom line dict[str(dominating result[0])].append(dominating result)
           off diag.append(board[self.board.BOARD SIDE-1-i, i])
       dominating result = self.dominating(diag, line counter)
       dominating result = self.dominating(off diag, line counter+1)
       if dominating result != None:
```

```
def dominating(self, line, line counter: int) -> list:
def check attributes(self) -> list:
```

```
if self.board.get piece charachteristics(elem).SQUARE:
piece = self.board.get piece charachteristics(i)
if piece.SOLID:
```

```
for i in range(len(dom line attributes)):
            desired piece[i].append(not(dom line attributes[i][0]))
def dominant attributes(self, dominant lines) -> list:
                dom line attributes[i].append(elem[2][i])
def find piece attribute(self, desired piece) -> tuple:
```

```
def place piece specified line(self, line) -> tuple([int, int]):
        line val.append(self.count line(row, piece))
        line val.append(self.count line(board[:,i], piece))
```

```
diag.append(board[i,i])
           off diag.append(board[self.board.BOARD SIDE-1-i, i])
       line val.append(self.count line(diag, piece))
       line val.append(self.count line(off diag, piece))
self.board.get piece charachteristics(piece).SQUARE:
       board = self.board.get board status()
           length.append(self.length of line(row))
           diag.append(board[i,i])
           off diag.append(board[self.board.BOARD SIDE-1-i, i])
```

```
line = list(dict.fromkeys(line))
            return len(line)
            return self.board.BOARD SIDE
        attribute values = self.check attributes()
self.place piece specified line(self.dom line dict['3'][0])
```

```
if call:
                    return self.place piece specified line(line)
list(self.piece dict[str(self.board.get selected piece())])):
                        return self.place piece specified line(line)
        piece = self.board.get selected piece()
        line values = self.count shared attributes in line(piece)
        return self.place piece specified line([0,
line values.index(max(line values))])
        return self.place piece specified line([0,
```

generate_agent_dumb.py:

```
import random
import quarto

class DumbAgent (quarto.Player):
    '''Evolved agent using the GA approach'''

def __init__ (self, quarto: quarto.Quarto) -> None:
    super().__init__ (quarto)
    self.board = quarto
    self.active_pieces = [i for i in range(16)]

    # Generate the probability of picking each rule for picking and
placing
    pick_prob_not_normalized = [random.random() for i in range(3)]
    pick prob = [x/sum(pick_prob_not_normalized) for x in
pick_prob_not_normalized]
    self.pick_probability_1 = pick_prob[0]
    self.pick_probability_2 = pick_prob[1]
    self.pick_probability_3 = pick_prob[2]

    place_probability_not_normalized = [random.random() for i in
range(6)]
    place_prob = [x/sum(place_probability_not_normalized) for x in
place_probability not_normalized]
    self.place_probability_1 = place_prob[0]
    self.place_probability_2 = place_prob[1]
    self.place_probability_3 = place_prob[2]
    self.place_probability_4 = place_prob[3]
    self.place_probability_5 = place_prob[5]
```

```
def get pick prob(self):
def get place prob(self):
    self.place probability 1 = val
def set place prob 6(self, val):
def choose piece(self) -> int:
```

```
self.dominating line()
```

```
def piece attribute dict(self) -> dict:
   def dominating line(self):
           dominating result = self.dominating(row, line counter)
self.dom line dict[str(dominating result[0])].append(dominating result)
       for i in range(self.board.BOARD SIDE):
self.dom line dict[str(dominating result[0])].append(dominating result)
       dominating result = self.dominating(diag, line counter)
       dominating result = self.dominating(off diag, line counter+1)
```

```
def dominating(self, line, line counter: int) -> list:
        piece = self.board.get piece charachteristics(elem)
def check attributes(self) -> list:
```

```
if self.board.get piece charachteristics(elem).SOLID:
        if self.board.get piece charachteristics(elem).SQUARE:
piece = self.board.get piece charachteristics(i)
   cur val += attribute values['Square']
```

```
desired piece[i].append(not(dom line attributes[i][0]))
def dominant attributes(self, dominant lines) -> list:
                dom line attributes[i].append(elem[2][i])
```

```
def place piece specified line(self, line) -> tuple([int, int]):
    board = self.board.get board status()
        line val.append(self.count line(board[:,i], piece))
```

```
off diag.append(board[self.board.BOARD SIDE-1-i, i])
line val.append(self.count line(diag, piece))
line val.append(self.count line(off diag, piece))
        if high == self.board.get piece charachteristics(piece).HIGH:
    length.append(self.length of line(row))
   diag.append(board[i,i])
   off diag.append(board[self.board.BOARD SIDE-1-i, i])
length.append(self.length of line(diag))
```

```
piece ranking = self.rank pieces(attribute values)
self.place piece specified line(self.dom line dict['3'][0])
```

```
call = True
piece = self.board.get selected piece()
line values = self.count shared attributes in line(piece)
return self.place piece specified line([0,
```

generate_agent_least_dumb.py:

```
import random
import quarto

class LeastDumbAgent(quarto.Player):
    '''Evolved agent using the GA approach'''

def __init__ (self, quarto: quarto.Quarto) -> None:
    super().__init__(quarto)
    self.board = quarto
    self.active_pieces = [i for i in range(16)]

    # Generate the probability of picking each rule for picking and

placing
    pick_prob_not_normalized = [random.random() for i in range(3)]
    pick_prob = [x/sum(pick_prob_not_normalized) for x in

pick_prob_not_normalized]
    self.pick_probability_1 = pick_prob[0]
    self.pick_probability_2 = pick_prob[1]
    self.pick_probability_3 = pick_prob[2]

    place_probability_not_normalized = [random.random() for i in

range(6)]
    place_probability_not_normalized = [random.random() for x in

place_probability_not_normalized]

    self.place_probability_1 = place_prob[0]
    self.place_probability_2 = place_prob[1]
    self.place_probability_3 = place_prob[2]
    self.place_probability_4 = place_prob[3]
    self.place_probability_5 = place_prob[5]

# Create a dictionary containing the attributes of each piece
```

```
def get pick prob(self):
def get place prob(self):
def set_place_prob_3(self, val):
def set place prob 6(self, val):
def choose piece(self) -> int:
```

```
piece dict[str(i)] = (piece.HIGH, piece.COLOURED, piece.SOLID,
```

```
piece.SQUARE)
           dominating result = self.dominating(row, line counter)
self.dom line dict[str(dominating result[0])].append(dominating result)
self.dom line dict[str(dominating result[0])].append(dominating result)
            off diag.append(board[self.board.BOARD SIDE-1-i, i])
        dominating result = self.dominating(diag, line counter)
   def dominating(self, line, line counter: int) -> list:
```

```
line = list(dict.fromkeys(line))
piece.SQUARE])
        attribue values['Solid'] = 0
        attribue values['Hollow'] = 0
```

```
if self.board.get piece charachteristics(elem).HIGH:
            attribue values['Circle'] += 1
if piece.SOLID:
if piece.SQUARE:
```

```
desired piece[i].append(not(dom line attributes[i][0]))
```

```
board = self.board.get board status()
    line val.append(self.count line(row, piece))
for i in range(self.board.BOARD SIDE):
    line val.append(self.count line(board[:,i], piece))
    diag.append(board[i,i])
    off diag.append(board[self.board.BOARD SIDE-1-i, i])
line val.append(self.count line(diag, piece))
line val.append(self.count line(off diag, piece))
```

```
self.board.get piece charachteristics(piece).SQUARE:
   def length of board lines(self) -> list:
            length.append(self.length of line(row))
        off diag = []
        for i in range(len(board)):
            diag.append(board[i,i])
            off diag.append(board[self.board.BOARD SIDE-1-i, i])
            return self.board.BOARD SIDE
```

```
piece ranking = self.rank pieces(attribute values)
self.place piece specified line(self.dom line dict['3'][0])
                    return self.place piece specified line(line)
```

```
def place rule 3(self):
        return self.place piece specified line([0,
        return self.place piece specified line([0,
lengths on board.index(min(lengths on board))])
        return self.place piece specified line([0,
lengths on board.index(max(lengths on board))])
```

```
'''Place the piece at random'''
board = self.board.get_board_status()

x = random.randint(0, 3)
y = random.randint(0, 3)

while board[x,y] != -1:
    x = random.randint(0, 3)
    y = random.randint(0, 3)
return (y, x)
```

generate_agent_less_dumb.py:

```
range(6)]
   def get_pick_prob(self):
```

```
def get place prob(self):
   self.dominating line()
```

```
def place piece(self) -> tuple([int, int]):
      self.dominating line()
                  return self.place piece specified line(line)
for i in range(16):
piece.SOUARE)
```

```
dominating result = self.dominating(row, line counter)
self.dom line dict[str(dominating result[0])].append(dominating result)
        for i in range(self.board.BOARD SIDE):
            dominating result = self.dominating(board[:,i], line counter)
self.dom line dict[str(dominating result[0])].append(dominating result)
            diag.append(board[i,i])
        dominating result = self.dominating(diag, line counter)
self.dom line dict[str(dominating result[0])].append(dominating result)
```

```
line = list(dict.fromkeys(line))
if len(line attributes) > 1:
    for i in range(len(line attributes)-1):
attribue values['Hollow'] = 0
attribue values['Square'] = 0
```

```
if self.board.get piece charachteristics(elem).HIGH:
        if self.board.get piece charachteristics(elem).SOLID:
   cur val += attribute values['Color']
if piece.SOLID:
```

```
desired_piece = [[], [], [], []]
        desired piece[i].append(not(dom line attributes[i][0]))
```

```
def place_piece_specified_line(self, line) -> tuple([int, int]):
   board = self.board.get board status()
    for i in range(self.board.BOARD SIDE):
        line val.append(self.count line(board[:,i], piece))
    for i in range(len(board)):
        off diag.append(board[self.board.BOARD SIDE-1-i, i])
    line val.append(self.count line(diag, piece))
    line val.append(self.count line(off diag, piece))
```

```
self.board.get piece charachteristics(piece).SQUARE:
            length.append(self.length of line(row))
        for i in range(len(board)):
            diag.append(board[i,i])
            off diag.append(board[self.board.BOARD SIDE-1-i, i])
        length.append(self.length of_line(diag))
```

```
max pos = piece ranking.index(max(piece ranking))
        while random piece in self.board.get board status():
self.place piece specified line(self.dom line dict['3'][0])
                if call:
```

```
return self.place piece specified line([0,
lengths on board.index(min(lengths on board))])
        return self.place piece specified line([0,
lengths on board.index(max(lengths on board))])
```

```
board = self.board.get_board_status()

x = random.randint(0, 3)

y = random.randint(0, 3)

while board[x,y] != -1:
    x = random.randint(0, 3)
    y = random.randint(0, 3)
return (y, x)
```

generate_individuals.py:

```
individuals and score.append([score, individual])
def gameplay(individual 1: EvolvedAgent, individual 2: EvolvedAgent, game:
```

```
def offspring(individuals: list, game: quarto.Quarto, mutation: int) -> list:
```

```
p1 = roulette(individuals)
           mutate(child1)
           mutate(child2)
       offspring.append(child1)
       offspring.append(child2)
def mutate(individual: EvolvedAgent):
def crossover(p1: EvolvedAgent, p2: EvolvedAgent, game: guarto.Quarto) ->
       child 1 pick.append(parent 1[i])
       child 2 pick.append(parent 2[i])
       child 1 pick.append(parent 2[i+pick cross])
```

```
parent 1 = p1.get_place_prob()
    child 1 place.append(parent 1[i])
    child 2 place.append(parent 2[i])
    child_1_place.append(parent_2[i+place_cross])
for in range(iterations):
   offspring individuals = offspring(individuals, game, mutate rate)
```

extendedquarto.py:

ga_agent.py:

```
import random
import quarto

class EvolvedAgent(quarto.Player):
    '''Evolved agent using the GA approach'''

def __init__(self, quarto: quarto.Quarto) -> None:
    super().__init__(quarto)
    self.board = quarto
    self.active_pieces = [i for i in range(16)]

    # Generate the probability of picking each rule for picking and
placing
    pick_prob_not_normalized = [random.random() for i in range(3)]
    pick_prob = [x/sum(pick_prob_not_normalized) for x in

pick_prob_not_normalized]
    self.pick_probability_1 = pick_prob[0]
    self.pick_probability_2 = pick_prob[1]
    self.pick_probability_3 = pick_prob[2]
```

```
place_probability_not_normalized = [random.random() for i in
    self.place probability 2 = place prob[1]
def set_pick_prob_2(self, val):
def set pick prob 3(self, val):
def set place prob 2(self, val):
def set place prob 3(self, val):
def choose piece(self) -> int:
```

```
self.place_probability_3 + self.place_probability 4:
piece.SQUARE)
            dominating result = self.dominating(row, line counter)
self.dom line dict[str(dominating result[0])].append(dominating result)
self.dom line dict[str(dominating result[0])].append(dominating result)
```

```
off diag.append(board[self.board.BOARD SIDE-1-i, i])
        dominating result = self.dominating(off diag, line counter+1)
    def dominating(self, line, line counter: int) -> list:
            line = list(dict.fromkeys(line))
            line attributes.append([piece.HIGH, piece.COLOURED, piece.SOLID,
piece.SQUARE])
                        result.append(None)
```

```
attribue values['Hollow'] += 1
for i in range(16):
```

```
if piece.SQUARE:
found, piece val = self.find piece attribute(desired piece)
for i in range(4):
```

```
dom line attributes[i].append(elem[2][i])
def find piece attribute(self, desired piece) -> tuple:
        if i not in self.board.get board status():
               and piece.SOLID in desired piece[2] and piece.SQUARE in
   board = self.board.get board status()
```

```
line val.append(self.count line(row, piece))
        for i in range(self.board.BOARD SIDE):
            line val.append(self.count line(board[:,i], piece))
            diag.append(board[i,i])
            off diag.append(board[self.board.BOARD SIDE-1-i, i])
        line_val.append(self.count_line(diag, piece))
        line val.append(self.count line(off diag, piece))
self.board.get piece charachteristics(piece).COLOURED:
self.board.get piece charachteristics(piece).SOLID:
self.board.get piece charachteristics(piece).SQUARE:
            length.append(self.length of line(row))
        for i in range(self.board.BOARD SIDE):
            length.append(self.length of line(board[:,i]))
```

```
diag.append(board[i,i])
    off diag.append(board[self.board.BOARD SIDE-1-i, i])
while min pos in self.board.get board status():
```

```
self.place piece specified line(self.dom line dict['3'][0])
                if call:
                        return self.place piece specified line(line)
        return self.place piece specified line([0,
line values.index(max(line values))])
```

```
return self.place piece specified line([0,
    return self.place piece specified line([0,
pick = [prob/sum(pick prob) for prob in pick prob]
place = [prob/sum(place prob) for prob in place prob]
```

```
plus_pick_025 = (0.06369372715277863, 0.8847749515760825,
0.051531321271138734)
plus_place_025 = (0.5117837231872915, 3.563965752635712e-05,
0.00032016772834300385, 0.07394996224198032, 0.38933362248027764,
0.024576884704581324)

set_pick_and_place(plus_agent_025, plus_pick_025, plus_place_025)
'''
```

ga_dumb.py:

```
self.place_probability_2 = place_prob[1]
self.place probability 4 = place prob[3]
```

```
self.place probability 3 = val
self.place probability 5 = val
```

```
self.dominating line()
```

```
dominating result = self.dominating(row, line counter)
           dominating result = self.dominating(board[:,i], line counter)
       dominating result = self.dominating(diag, line counter)
self.dom line dict[str(dominating result[0])].append(dominating result)
       dominating result = self.dominating(off diag, line counter+1)
   def dominating(self, line, line counter: int) -> list:
       line attributes = []
```

```
result.append(j)
def check attributes(self) -> list:
    attribue values['High'] = 0
    attribue values['Circle'] = 0
```

```
if self.board.get piece charachteristics(elem).HIGH:
        if self.board.get piece charachteristics(elem).COLOURED:
        if self.board.get piece charachteristics (elem) . SQUARE:
   cur val += attribute values['Noncolor']
if piece.SOLID:
```

```
desired piece[i].append(True)
                desired piece[i].append(False)
dom line attributes[i]:
    def find piece attribute(self, desired piece) -> tuple:
    def place piece specified line(self, line) -> tuple([int, int]):
```

```
the agent wins''
           line val.append(self.count line(board[:,i], piece))
       line val.append(self.count line(diag, piece))
       line val.append(self.count line(off diag, piece))
```

```
line = list(dict.fromkeys(line))
def length_of_board_lines(self) -> list:
        length.append(self.length_of_line(row))
        length.append(self.length of line(board[:,i]))
        off diag.append(board[self.board.BOARD SIDE-1-i, i])
    length.append(self.length of line(diag))
    length.append(self.length of line(off diag))
        return self.board.BOARD SIDE
```

```
call = True
    return self.place piece specified line(line)
```

```
def place rule 2(self):
    return self.place piece specified line([0,
```

```
x = random.randint(0, 3)
y = random.randint(0, 3)

while board[x,y] != -1:
    x = random.randint(0, 3)
    y = random.randint(0, 3)
return (y, x)
```

ga_less_dumb.py:

```
def init (self, quarto: quarto.Quarto) -> None:
    self.pick probability 2 = pick prob[1]
```

```
def set_pick_prob_2(self, val):
def set_place prob 6(self, val):
    self.place probability 6 = val
```

```
self.place probability 3:
piece.SQUARE)
    def dominating line(self):
```

```
board = self.board.get board status()
           dominating result = self.dominating(row, line counter)
           dominating result = self.dominating(board[:,i], line counter)
self.dom line dict[str(dominating result[0])].append(dominating result)
           diag.append(board[i,i])
           off diag.append(board[self.board.BOARD SIDE-1-i, i])
       dominating result = self.dominating(diag, line counter)
       dominating result = self.dominating(off diag, line counter+1)
self.dom line dict[str(dominating result[0])].append(dominating result)
   def dominating(self, line, line counter: int) -> list:
```

```
piece = self.board.get piece charachteristics(elem)
                    result.append(None)
def check attributes(self) -> list:
```

```
cur val += attribute values['Solid']
if piece.SQUARE:
val.append(cur val)
```

```
dom line attributes[i].append(elem[2][i])
    return dom line attributes
    for i in range(16):
                and piece.SOLID in desired piece[2] and piece.SQUARE in
def place piece specified line(self, line) -> tuple([int, int]):
```

```
line val.append(self.count line(row, piece))
for i in range(self.board.BOARD SIDE):
off diag = []
    diag.append(board[i,i])
```

```
high, color, solid, square = self.piece dict[str(elem)]
def length of board lines(self) -> list:
    board = self.board.get board status()
        length.append(self.length of line(board[:,i]))
    off diag = []
    length.append(self.length of line(off diag))
    piece ranking = self.rank pieces(attribute values)
```

```
min pos = piece ranking.index(min(piece ranking))
self.place piece specified line(self.dom line dict['3'][0])
                        call = False
                    return self.place piece specified line(line)
   def place rule 2(self):
```

```
'''Place the piece on the longest uninterrupted line with a shared
return self.place piece specified line([0,
return self.place piece specified line([0,
board = self.board.get board status()
```

```
while board[x,y] != -1:
    x = random.randint(0, 3)
    y = random.randint(0, 3)

return (y, x)
```

main.py:

```
game = quarto.Quarto()
random agent = RandomPlayer(game)
logging.warning(f"main: Winner: player {winner}")
    logging.getLogger().setLevel(level=logging.WARNING)
```

```
elif args.verbose == 1:
    logging.getLogger().setLevel(level=logging.INFO)
elif args.verbose == 2:
    logging.getLogger().setLevel(level=logging.DEBUG)
main()
```

Opponent_Agents.py:

```
from quarto import Player
 def choose piece(self) -> int:
```

rl.py:

```
import numpy as np
from testQuarto import TestQuarto
       return hash((hash(self.board.tostring()), self.selected piece))
   def get board str(self):
   def get_selected_piece str(self):
```

```
PENALTY = -1 #Penalty value for losing the game
def clear_previous vars(self) -> None:
def get q length(self) -> int:
def generate possible moves(self) -> list:
```

```
board = current state.get board status()
              not selected pieces.remove(index at place)
               possible moves.append((empty place[0], empty place[1],
def add new state move(self) -> None:
       current key = QTableKey(board, selected piece) #creates key
```

```
def policy(self) -> tuple:
def update when draw(self) -> None:
```

```
def update q(self) -> tuple:
        if (next state.check finished() and (next state.check winner() ==
```

```
for key, value in self.q.items():
   f.write(elem)
```

testQuarto.py:

```
return self._board
```

train_Q_learner:

```
player = 0
    q learner player = 0
   players = (external agent, q learner)
        piece ok = game.select(players[player].choose piece())
        game.set current player(1)
```

```
def q_learning_strategy(game: extendedQuarto.ExtendedQuarto, q_learner:
RLPlayer, num games: int):
       game.set players((opponent, q learner))
```

```
winner = train_q_learner(game, q_learner, opponent,
        results.append("won")
print("Lost: " + str(lost))
print("Draw: " + str(draw))
```

```
game_train = extendedQuarto.ExtendedQuarto()
    q_learner = RLPlayer(game_train, learning_rate, discount_rate,
exploration_rate)

    q_learning_strategy(game_train, q_learner, 3000)

    q_learner.save_q_table() #saves table after training q-learner
    # q_table_1: num_games = 1000, games added per opponent = 500,
exploration rate = decreases by 0,05 every 100th game
    # q_table_2: num_games = 2000, games added per opponent = 700,
exploration rate = decreases by 0,05 every 200th game
    # q_table_3: num_games = 3000, games added per opponent = 700,
exploration rate = decreases by 0,05 every 300th game

run()
```

trained_rl:

```
from quarto import Player
import numpy as np
class TrainedRL(Player):
 def choose piece(self) -> int:
```

```
given piece = int(elems[1])
 key = QTableKey(board, given piece)
f.close()
           empty places.append((x,y))
            not selected pieces.remove(index at place)
            possible moves.append((empty place[0], empty place[1],
```

SOURCE CODE FOR LAB 1:

List generation and class State.py:

```
# 'Taking inspiration from the puzzle problem and professor's code, defining class 'State'
import logging
from random import seed, choice
from typing import Callable
from gx_utils import *
import random
def problem(N, seed=None):
  random.seed(seed)
  return [
    list(set(random.randint(0, N - 1) for n in range(random.randint(N // 5, N // 2))))
    for n in range(random.randint(N, N * 5))
  1
class State:
  def init (self, data: set):
    self._data = set(data)
  # 'hashing by Frozenset to avoid any mutation'
  def __hash__(self):
    return hash(frozenset(self._data))
  # since the length of lists will be important to find the solution
  def __len__(self):
    return len(self._data)
  def __eq__(self, other):
    return self._data == other._data
  def __lt__(self, other):
    return self._data < other._data
  def __or__(self, other):
```

```
return State(self._data | other._data)
  def __and__(self, other):
    return State(self._data & other._data)
  def __sub__(self, other):
    return State(self._data - other._data)
  def __str__(self):
    return str(self._data)
  def __repr__(self):
    return repr(self._data)
  def data(self):
    return self._data
  def copy_data(self):
    return self._data.copy()
# 'Writing Search algorithm for subsequent searches to find the final solution'
def search(
    initial_state: State,
    goal_test: Callable,
    parent_state: dict,
    state_cost: dict,
    priority_function: Callable,
    unit_cost: Callable,
):
  frontier = PriorityQueue()
  parent_state.clear()
  state_cost.clear()
  state = initial_state
  parent_state[state] = None
  state_cost[state] = 0
  while state is not None and not goal_test(state):
```

```
for a in possible_actions():
      new_state = result(state, a)
      cost = unit_cost(a)
      if new_state not in state_cost and new_state not in frontier:
        parent_state[new_state] = state
        state_cost[new_state] = state_cost[state] + cost
        frontier.push(new_state, p=priority_function(new_state))
        logging.debug(f"Added new node to frontier (cost={state_cost[new_state]})")
      elif new_state in frontier and state_cost[new_state] > state_cost[state] + cost:
        old_cost = state_cost[new_state]
        parent_state[new_state] = state
        state cost[new state] = state cost[state] + cost
        logging.debug(f"Updated node cost in frontier: {old cost} ->
{state_cost[new_state]}")
    if frontier:
      state = frontier.pop()
    else:
      state = None
  path = list()
  s = state
  while s:
    path.append(s.copy_data())
    s = parent_state[s]
  logging.info(f"Found a solution in {len(path);,} steps; The number of visited states:
{len(state_cost):,} ")
  return list(reversed(path))
def goal_test(state):
  return state == GOAL
def possible_actions():
  return (State(x) for x in P)
def result(state, action):
  return state | action
```

```
# 'Breadth First'
logging.getLogger().setLevel(logging.INFO)
for N in [5, 10, 20, 100, 500, 1000]:
  parent_state = dict()
  state_cost = dict()
  GOAL = State(set(range(N)))
  P = problem(N, seed=42)
  def h(state):
    return len(state)
  INITIAL_STATE = State(set())
  logging.info(f'N = {N}')
  final = search(
    INITIAL_STATE,
    goal_test=goal_test,
    parent_state=parent_state,
    state_cost=state_cost,
    priority_function=lambda s: h(s),
    unit_cost=lambda a: 1,
  )
  logging.debug(final)
# 'DEPTH FIRST
logging.getLogger().setLevel(logging.INFO)
for N in [5, 10, 20, 100, 500, 1000]:
  parent_state = dict()
  state_cost = dict()
  GOAL = State(set(range(N)))
  P = problem(N, seed=42)
  INITIAL_STATE = State(set())
  logging.info(f'N = {N}')
  final = search(
```

```
INITIAL_STATE,
    goal_test=goal_test,
    parent_state=parent_state,
    state_cost=state_cost,
    priority_function=lambda s: -len(state_cost),
    unit_cost=lambda a:1
  )
  logging.debug(final)
import heapq
from collections import Counter
class PriorityQueue:
  """A basic Priority Queue with simple performance optimizations"""
  def __init__(self):
    self._data_heap = list()
    self._data_set = set()
  def __bool__(self):
    return bool(self._data_set)
  def __contains__(self, item):
    return item in self._data_set
  def _len_(self):
    return len(self._data_set)
  def push(self, item, p=None):
    assert item not in self, f"Duplicated element"
    if p is None:
      p = len(self._data_set)
    self._data_set.add(item)
    heapq.heappush(self._data_heap, (p, item))
  def pop(self):
    p, item = heapq.heappop(self._data_heap)
    self._data_set.remove(item)
    return item
```

```
class Multiset:
  """Multiset"""
  def __init__(self, init=None):
    self._data = Counter()
    if init:
      for item in init:
        self.add(item)
  def __contains__(self, item):
    return item in self._data and self._data[item] > 0
  def __getitem__(self, item):
    return self.count(item)
  def __iter__(self):
    return (i for i in sorted(self._data.keys()) for _ in range(self._data[i]))
  def _len_(self):
    return sum(self._data.values())
  def __copy__(self):
    t = Multiset()
    t._data = self._data.copy()
    return t
  def __str__(self):
    return f"M{{{', '.join(repr(i) for i in self)}}}"
  def __repr__(self):
    return str(self)
  def __or__(self, other: "Multiset"):
    tmp = Multiset()
    for i in set(self._data.keys()) | set(other._data.keys()):
      tmp.add(i, cnt=max(self[i], other[i]))
    return tmp
```

```
def __and__(self, other: "Multiset"):
  return self.intersection(other)
def __add__(self, other: "Multiset"):
  return self.union(other)
def __sub__(self, other: "Multiset"):
  tmp = Multiset(self)
  for i, n in other._data.items():
    tmp.remove(i, cnt=n)
  return tmp
def __eq__(self, other: "Multiset"):
  return list(self) == list(other)
def __le__(self, other: "Multiset"):
  for i, n in self._data.items():
    if other.count(i) < n:</pre>
      return False
  return True
def __lt__(self, other: "Multiset"):
  return self <= other and not self == other
def __ge__(self, other: "Multiset"):
  return other <= self
def __gt__(self, other: "Multiset"):
  return other < self
def add(self, item, *, cnt=1):
  assert cnt >= 0, "Can't add a negative number of elements"
  if cnt > 0:
    self._data[item] += cnt
def remove(self, item, *, cnt=1):
  assert item in self, f"Item not in collection"
  self._data[item] -= cnt
  if self._data[item] <= 0:</pre>
    del self._data[item]
```

```
def count(self, item):
    return self._data[item] if item in self._data else 0

def union(self, other: "Multiset"):
    t = Multiset(self)
    for i in other._data.keys():
        t.add(i, cnt=other[i])
    return t

def intersection(self, other: "Multiset"):
    t = Multiset()
    for i in self._data.keys():
        t.add(i, cnt=min(self[i], other[i]))
    return t
```