**PACMAN ASSIGNMENT ON CLASSICAL SEARCH**

QUESTION 1:

1. **Is this implementation recursive?**

Yes, the implementation is recursive. Because in this algorithm, we started with a root node and pushed to stack and checked if the node popped from stack is goal state. If yes, we stopped, else we continued with checking the currently processed node’s neighbours and again inserted it into stack. Then the same cycle is continued for each node till goal node is found. Hence it is a recursive algorithm.

1. **In building the graph, what information does each node contain?**

Each node contains 3 information:

1. State
2. Action
3. Cost
4. **What is the data structure used in this implementation of DFS?**

Stack data structure is used in DFS.

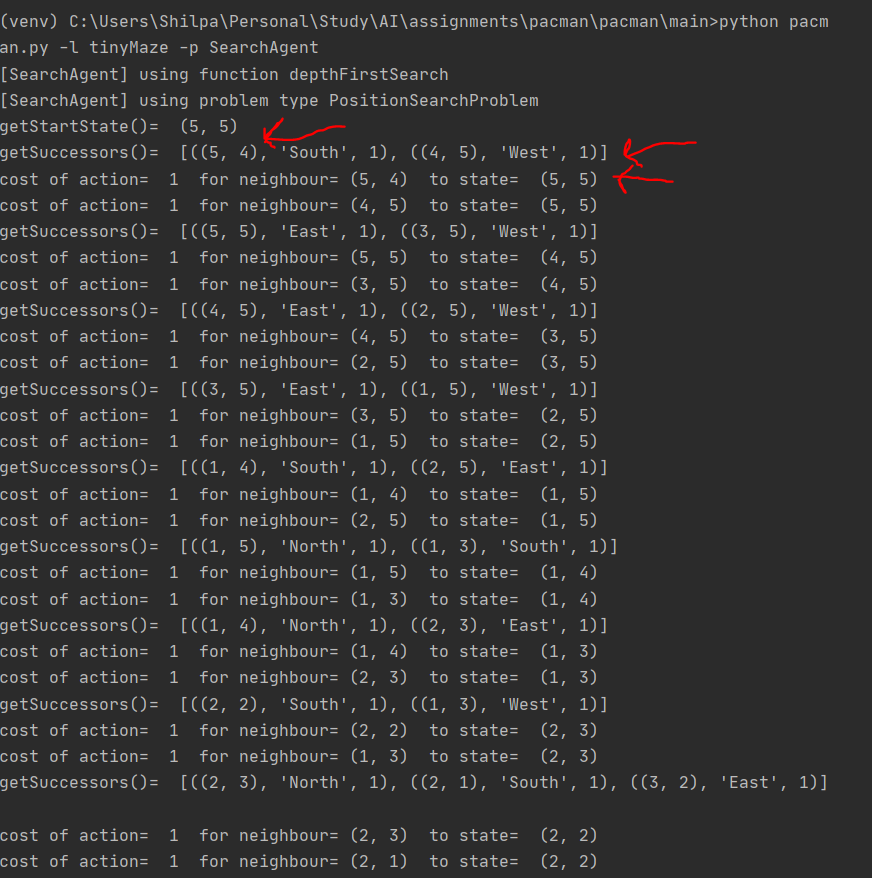
1. **In utils.py, you are provided with additional implemented data structures useful for completing the next questions. Provide a list of them and briefly explain the pros and cons of each one.**

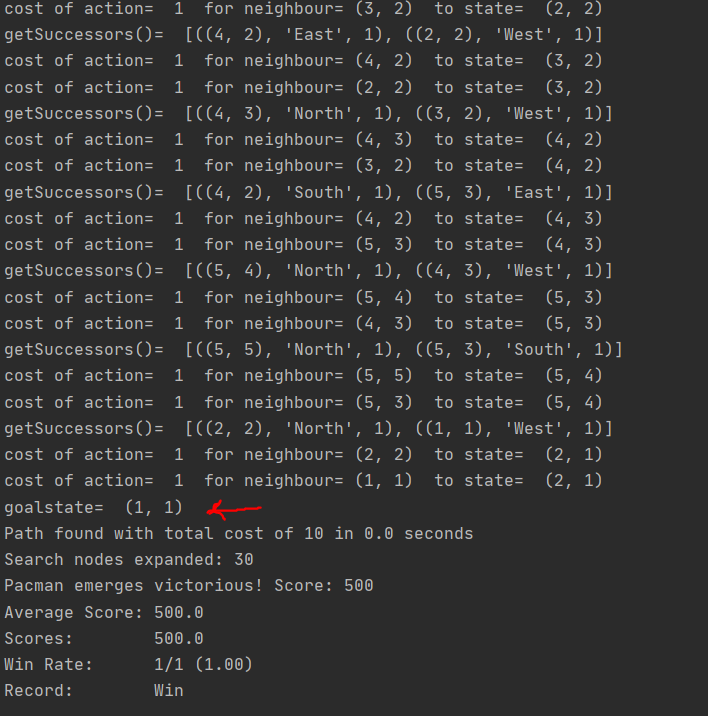
|  |  |  |  |
| --- | --- | --- | --- |
| Data Structure mentioned in util.py | Pros | Cons | Use |
| Stack | Stack is a LIFO(Last-In-First-Out) Data structure. If we want to check along the successors of immediately processed node, Stack is useful because we can use the latest processed nodes Hence its space complexity is relatively lesser when we use stack | However if the depth of the graph reaches to an infinity, the algorithms using Stack (e.g. DFS) gets lost or go on for infinite amount of time. The time complexity is huge and there is no guarantee that any solution state can be found, it might just get stuck in a branch of infinite depth. He | DFS(Depth First Search) |
| Queue | Queue helps us in level order searching of a graph, it is FIFO(First-In-First-Out) data structure.  If we want to search for an element it will take O(n) time. The advantage of using queue is that all nodes of a tree can be for sure covered to find a solution | Since queue uses FIFO while processing, it might be possible that the algorithm using it can create a huge queue of items to process till the earlier items get processed. Hence it requires huge amount of memory space | BFS(Breadth First Search) |
| Priority Queue | It helps in storing the root element based on some priority(min heap or max heap) The advantage is that retrieval of element is in O(1) | Disadvantage is that it take O(logn) to store the element at correct place using Heapify. Hence everytime we eant to take out k-elements. K-heapify actions would be needed | Uniform Cost Search |

1. **As mentioned earlier, the abstract SearchProblem class has four methods. Try to see the functionality of each by printing the outputs.**

The four methods are:

1. getStartState: returns the state state position (x,y)
2. isGoalState: returns the goal state position (x,y)
3. getSuccessors: get the successor lists. Each item has 3 values: position of successor in (x,y), Direction wrt the current node and the cost from current state to the particular succeccore node.
4. getCostOfActions: returns the cost of action to process one node to get its corresponsing successor.



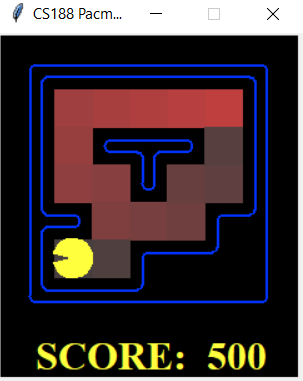
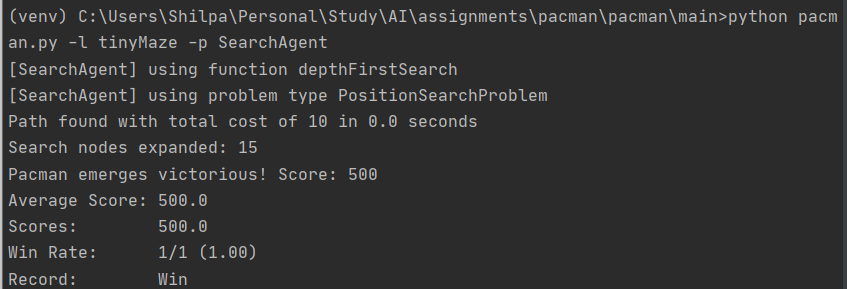


1. **Given the above information, explain the working flow of the implemented DFS**
2. Get the problem start state store it at root node, assigning action as empty initially and cost as 0 to begin with.
3. Push this root node to stack
4. While stack is not empty, pop the top element.
5. If the element’s state denotes the goal state, terminate the algorithm.
6. Else update the visited tag of the popped element and get its one neighbour, process it(update the action and cost) and push it in stack.
7. Step c-d is continued till stack is empty.

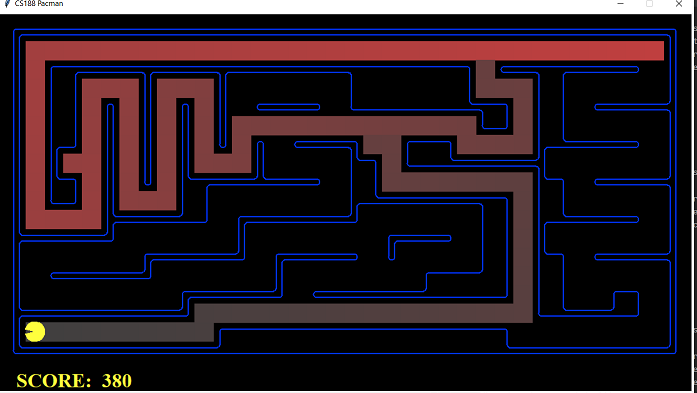
Basically DFS select one particular and goes down its depth, processes all nodes and then moves to the nearest parent’s other branch and continues the process till goal state is found.

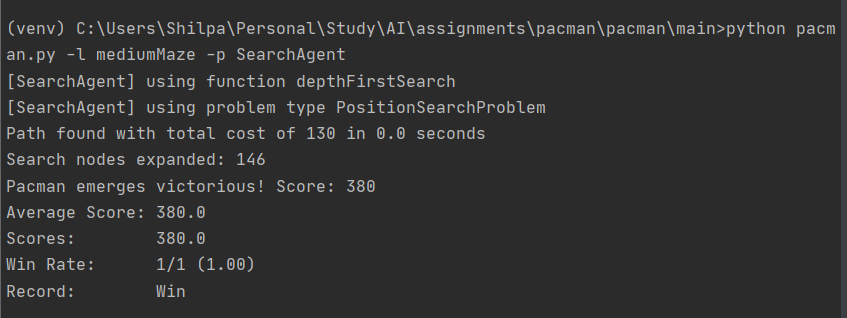
1. **Execute the following commands code should quickly find a solution:**

python pacman.py -l tinyMaze -p SearchAgent

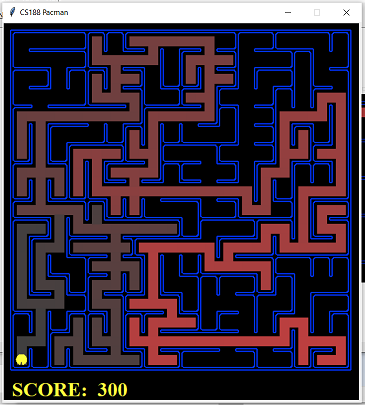
 

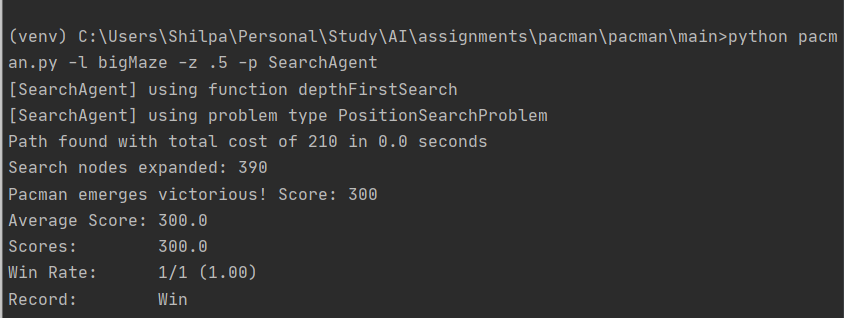
python pacman.py -l mediumMaze -p SearchAgent





python pacman.py -l bigMaze -z .5 -p SearchAgent

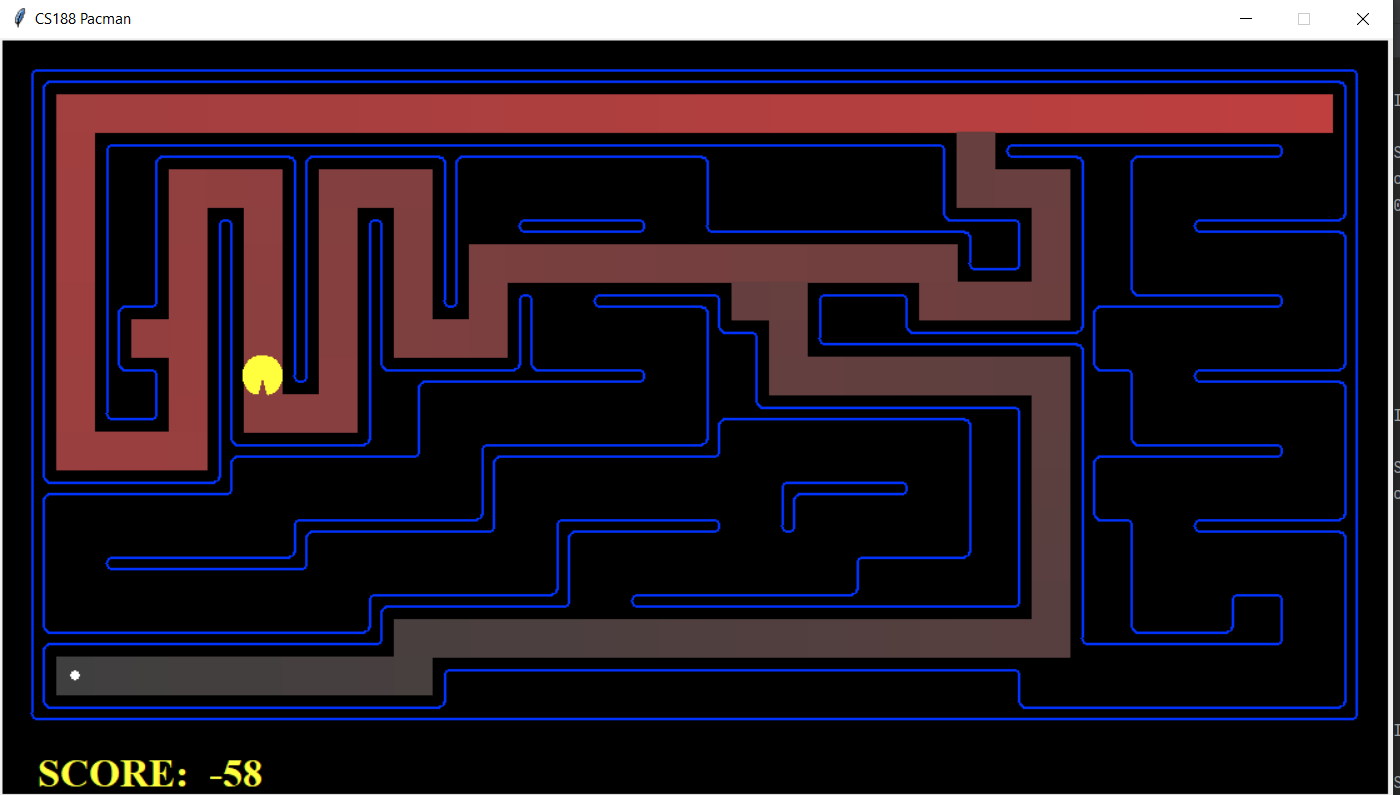




1. **The Pacman board will show an overlay of the states explored, and the order in which they were explored (brighter red means earlier exploration). Is the exploration order what you would have expected? Does Pacman actually go to all the explored squares on his way to the goal?**

Yes, since this is a DFS algorithm the Pacman is taking the first branch it is finding. Its along the depth.

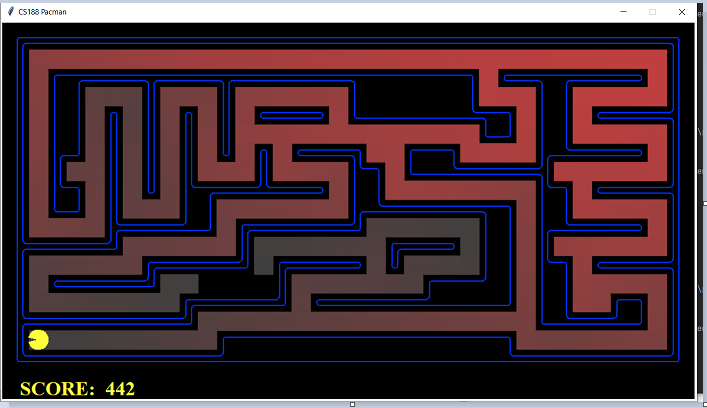
However, to answer the second question, it does not go to all explored squares on its way to the goal. Since it is a Depth First Search, it might be possible that one branch is explored completetly and goal is not found, then DFS goes to the nearest parent with an unexplored child and starts from there. Because of this it might be possible that the nodes have been explored , but they have not been considered for the goal path.

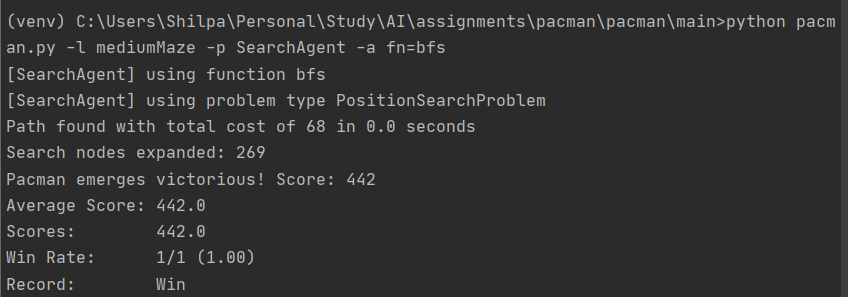


**QUESTION 2:**

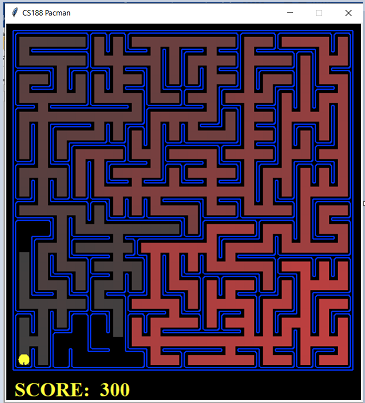
**Test your code the same way you did for depth-first search.**

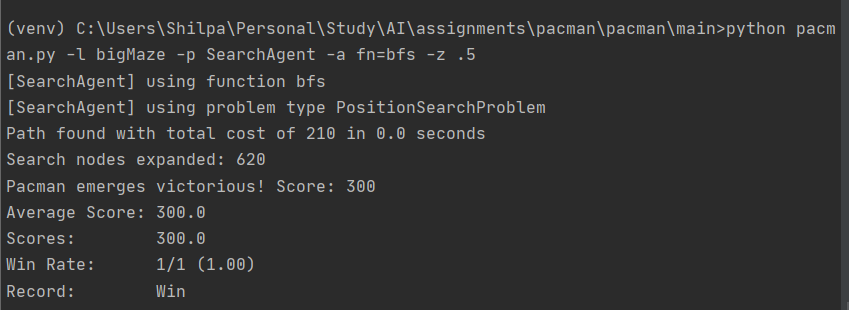
1. **python pacman.py -l mediumMaze -p SearchAgent -a fn=bfs**





1. **python pacman.py -l bigMaze -p SearchAgent -a fn=bfs -z .5**





**Does BFS find a least cost solution? If not, check your implementation**

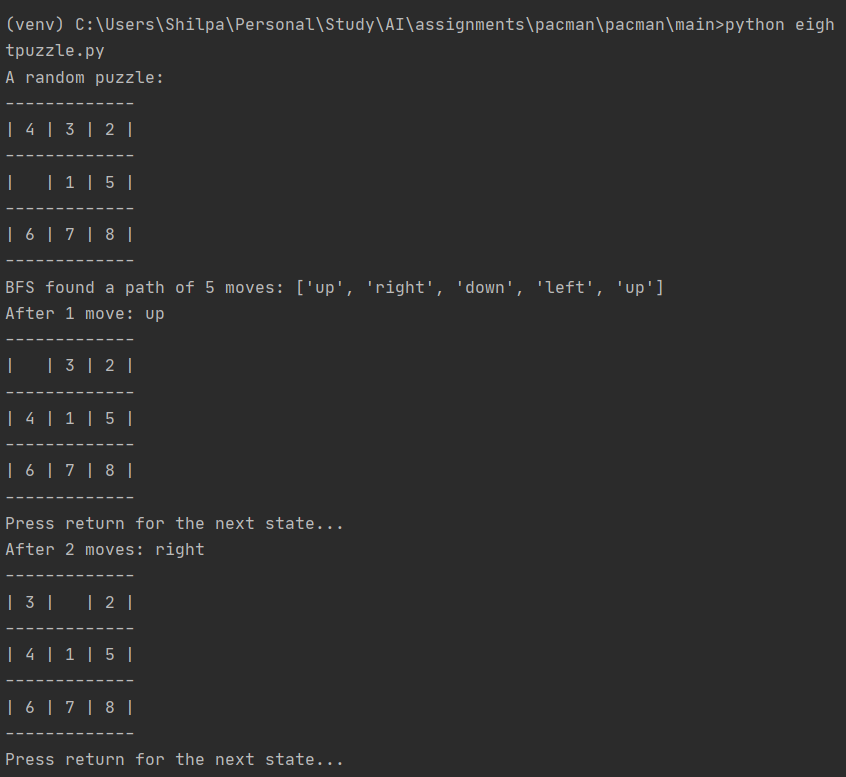
**Observation:**

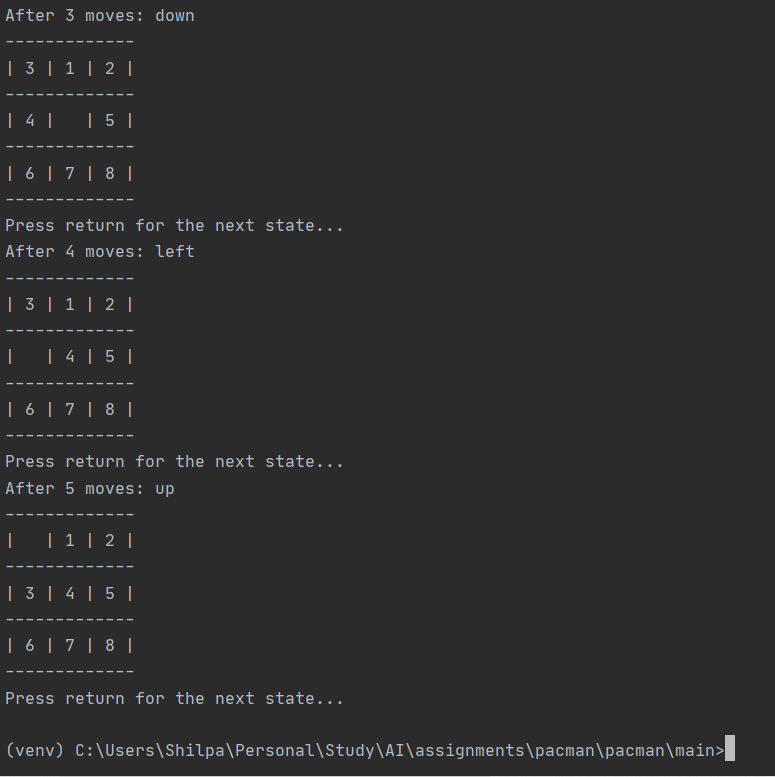
Yes. Here are the given comparison from above screenshots:

|  |  |  |
| --- | --- | --- |
| Algorithm | Mazetype | Cost |
| BFS | Medium | 68 |
| BFS | Large | 210 |
| DFS | Medium | 130 |
| DFS | Large | 210 |

**Observation:** In both the cases we see that the cost of BFS search<=cost of DFS Search

**If you've written your search code generically, your code should work equally well for the eight-puzzle search problem without any changes. python eightpuzzle.p**





**QUESTION 3:**

**Observation:**

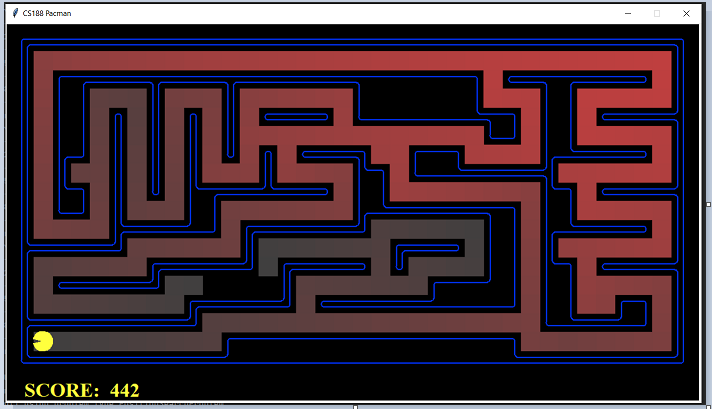
1. As per question it is observed that very low and very high path costs exist for the StayEastSearchAgent and StayWestSearchAgent respectively, due to their exponential cost functions . Below is the given comparison table and the screenshots of the execution along with commands:

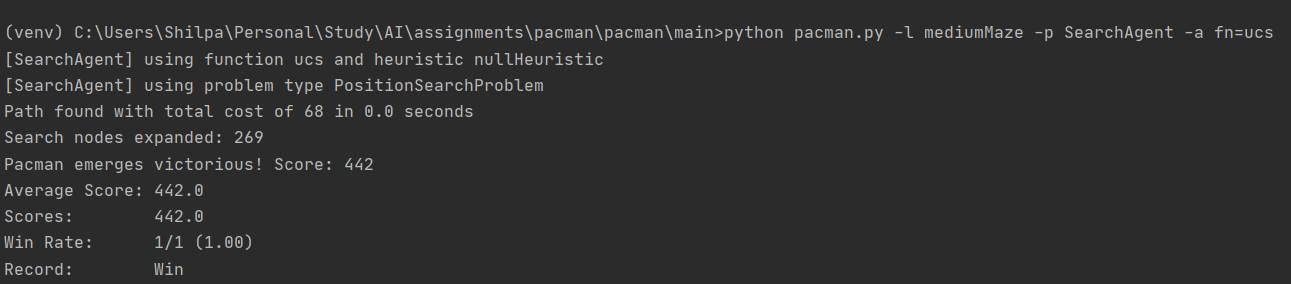
|  |  |
| --- | --- |
| **Maze path** | **Cost** |
| StayEastSearchAgent | 1 |
| StayWestSearchAgent | 68719479864 |
| Basic UCS | 68 |

1. Observed successful behaviours in all three of the following layout. The difference in cost is compared in above table.

**Basic UCS run on mediumMaze:**

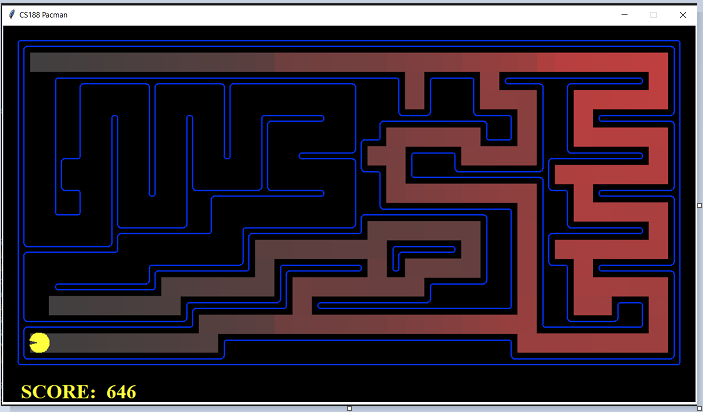
1. **python pacman.py -l mediumMaze -p SearchAgent -a fn=ucs**

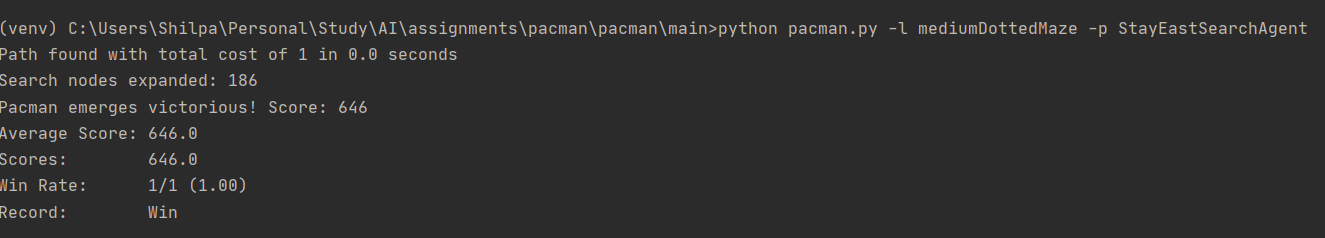




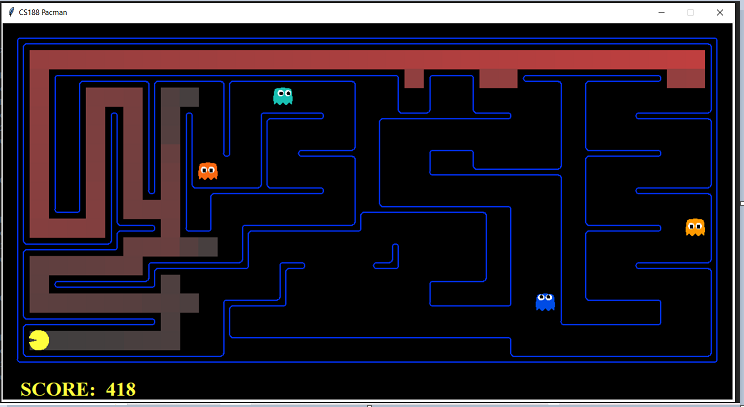
**East and West Agent runs:**

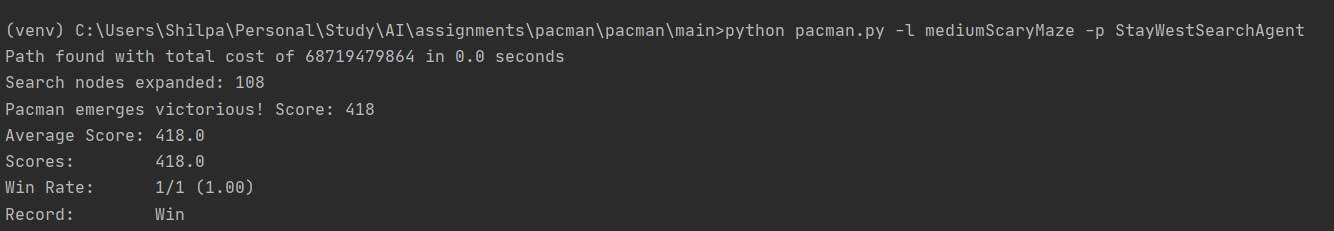
1. **python pacman.py -l mediumDottedMaze -p StayEastSearchAgent**





1. **python pacman.py -l mediumScaryMaze -p StayWestSearchAgent**

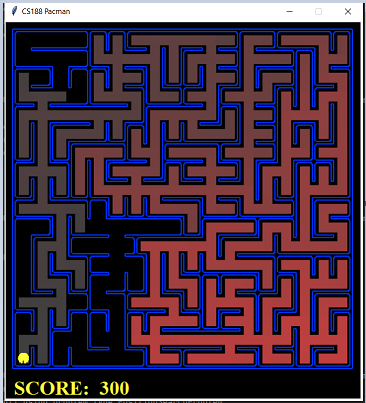


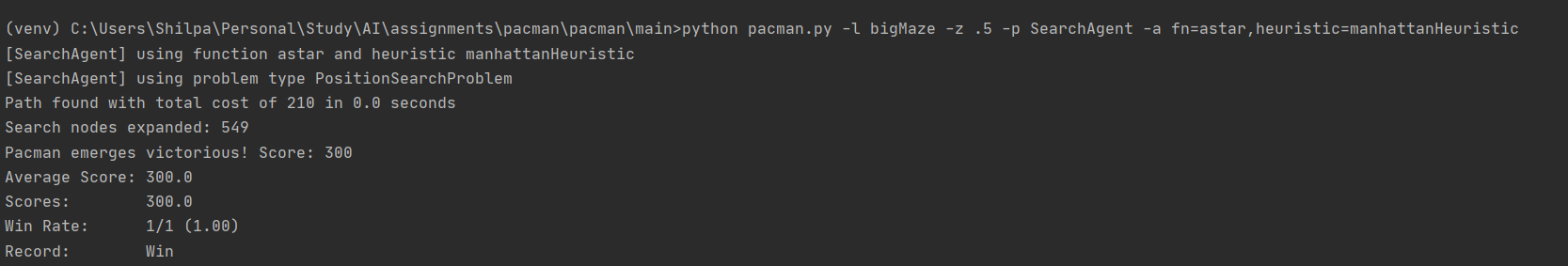


**QUESTION 4:**

**Basic implementation testing:**

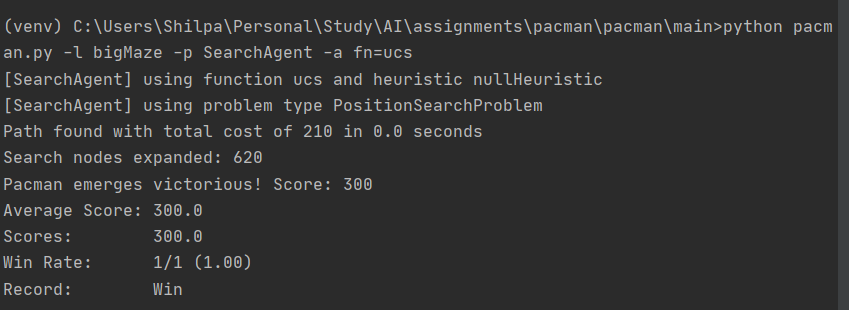
1. **python pacman.py -l bigMaze -z .5 -p SearchAgent -a fn=astar,heuristic=manhattanHeuristic**



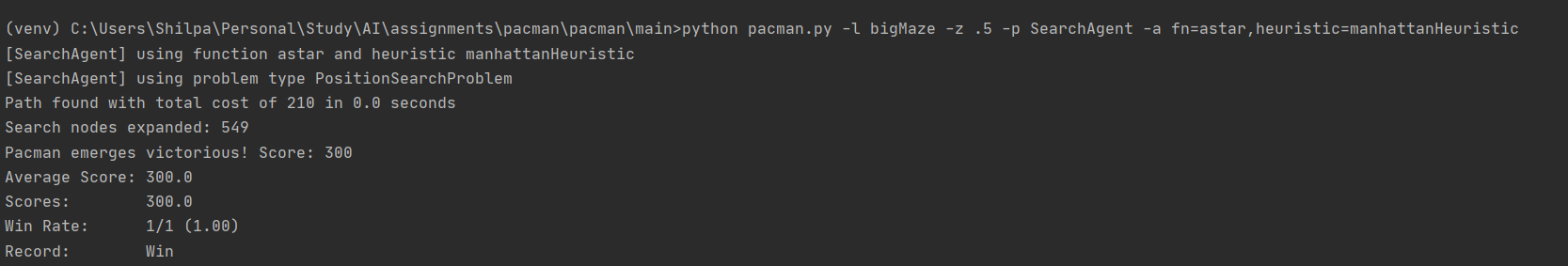


**Observation as per question:** A\* finds the optimal solution slightly faster than uniform cost search (about 549 vs. 620 search nodes expanded in our implementation, but ties in priority may make your numbers differ slightly).

**A-**star



**UCS**



|  |  |  |
| --- | --- | --- |
|  | **UCS** | **A-Star** |
| **Nodes expanded** | **620** | **549** |

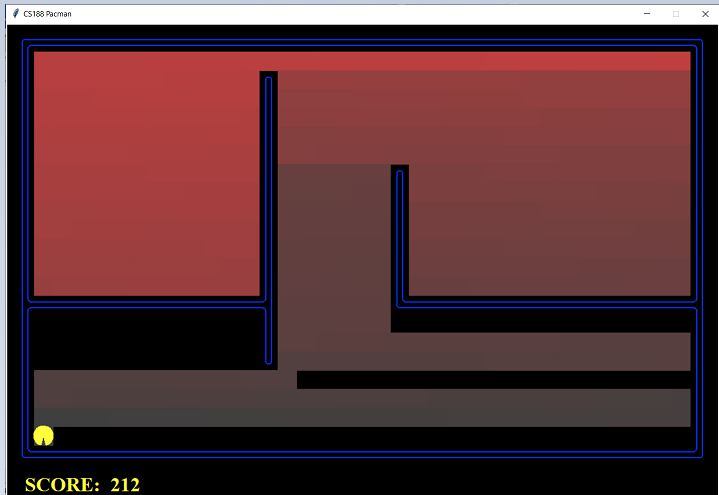
**What happens on openMaze for the various search strategies**

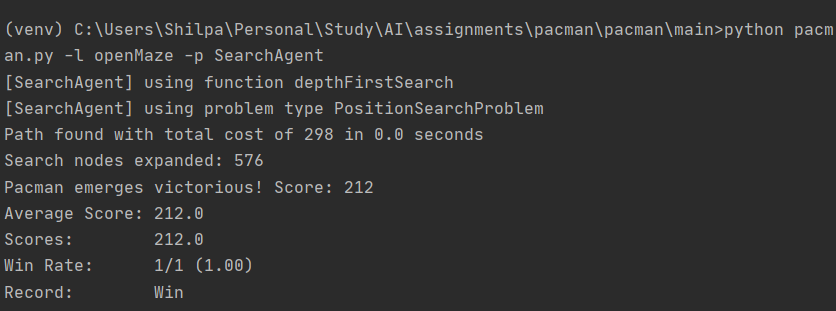
Comparison for all 4 algorithms on openMaze:

|  |  |  |  |
| --- | --- | --- | --- |
| Algorithms | Cost | No of Nodes Expanded | Score |
| DFS | 298 | 576 | 212 |
| BFS | 54 | 682 | 456 |
| UCS | 54 | 682 | 456 |
| A-Star | 54 | 535 | 456 |

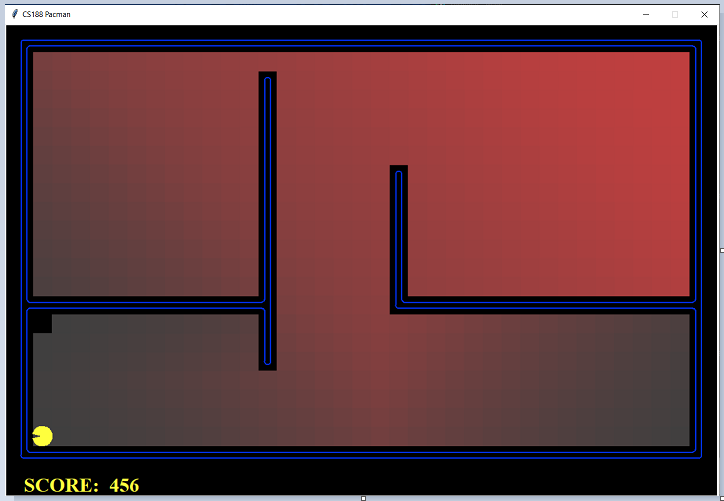
Observation: A-Star performs the best of all with minimum number of nodes expanded, lease cost. BFS and UCS has same performance. DFS is not an optimal algorithm , hence its cost is more than the others.

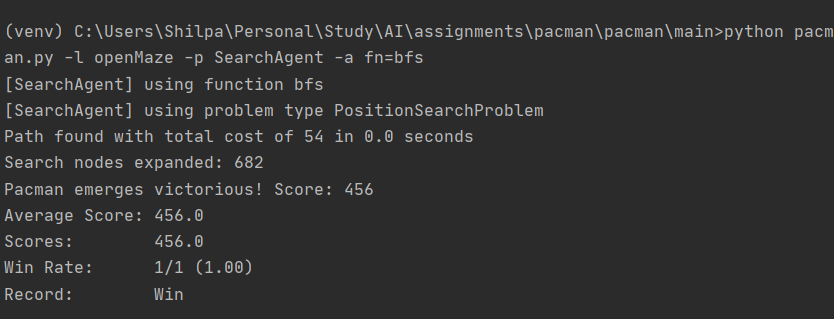
1. **DFS:** python pacman.py -l openMaze -p SearchAgent



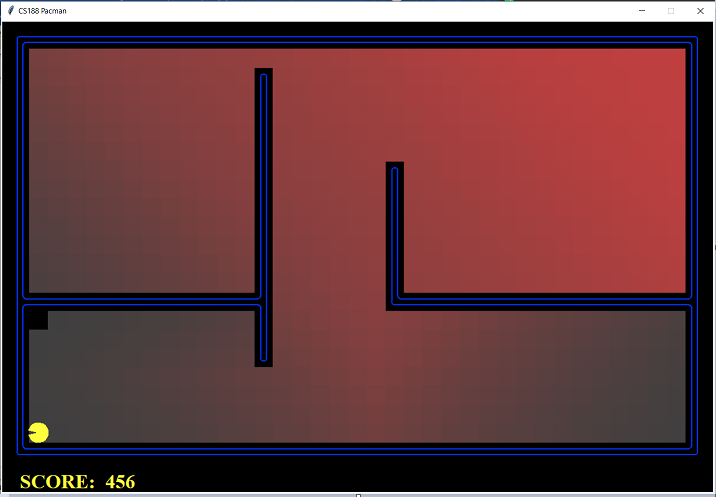


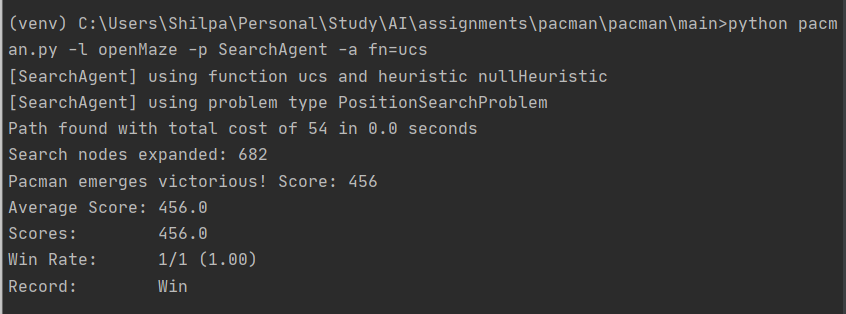
1. **BFS:** python pacman.py -l openMaze -p SearchAgent -a fn=bfs



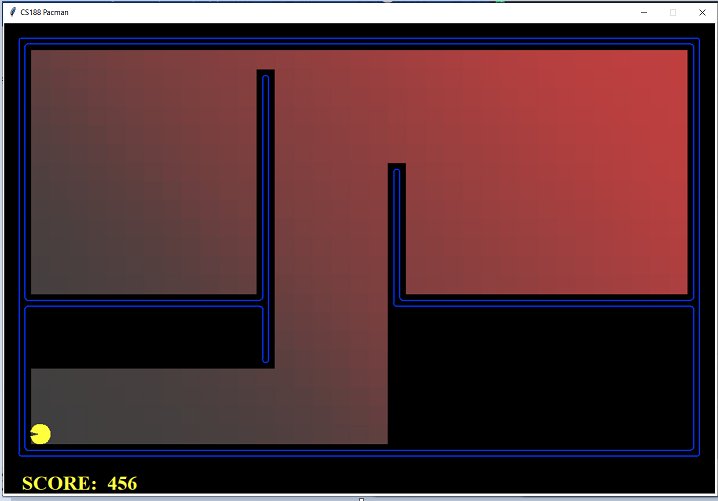


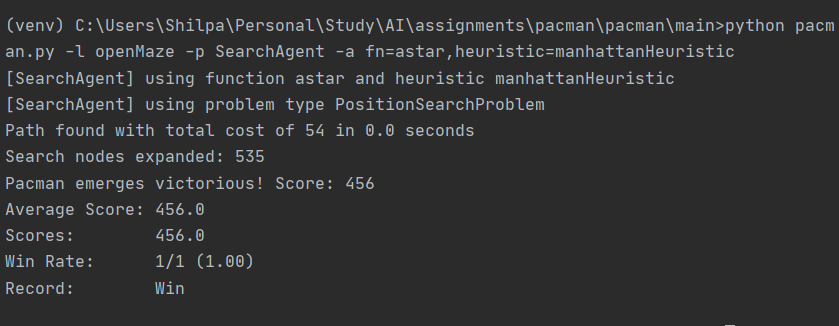
1. **UCS:** python pacman.py -l openMaze -p SearchAgent -a fn=ucs





1. **A-star:** python pacman.py -l openMaze -p SearchAgent -a fn=astar,heuristic=manhattanHeuristic





**QUESTION 5:**

1. **What is the state representation? In other words, what information are encoded for each state?**
2. Current location of the agent (x,y)
3. Details of corners unexplored as list storing the (x,y) coordinate of each corner.

Eg. In the below screenshot, once the agent reaches(6,6) which an unexplored corner enlisted in the corner list, that particular corner is removed from list.s



1. **An abstract state representation should not encode irrelevant information (like the position of ghosts, where extra food is, etc.). In particular, using Pacman GameState as a search state should be avoided or your code will be very, very slow if you do (and also wrong). Can you think of alternative state representation satisfying these properties?**

There can be many alternate representations Two kinds of alternate representation that I can think are noted down below and a little discussion about both:

1. State= ((position x,y), (Manhattan Distance to the nearest Corner), (Direction to that nearest corner))

We can have a separate list of corner positions, we can keep on checking if any of the goal state item is reached and remove them. The final goal state will be an empty list of corners remaining.

1. State= ((position x,y), (List of available Actions from the current state))

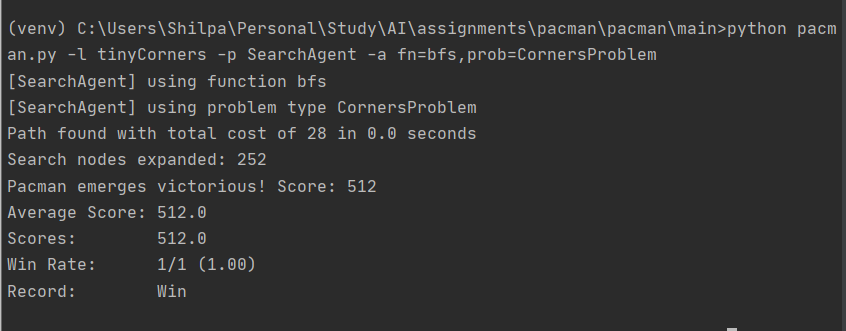
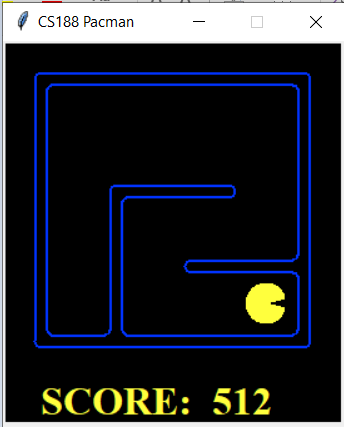
In this case, we can check the heuristic value (e.g. Manhattan Distance or Euclidean Distance) of the next state after each item in the action list is completed. Then the action which leads to a state with minimum heuristic value can be taken up.

1. **How does the agent check whether a state is a goal state?**

In order for a state to be a goal state, the number of corners explored has to be zero. Hence the length of remaining number of corner=0

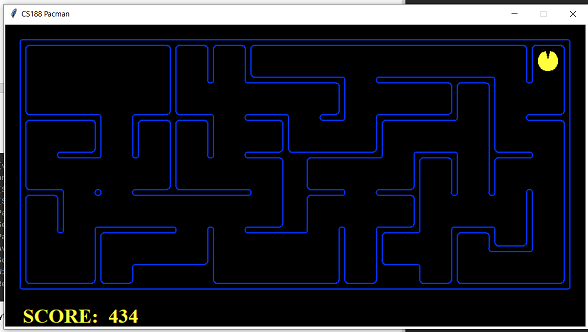
4. **Implement getSuccessors(self, state) Now, your search agent should be able to solve:**

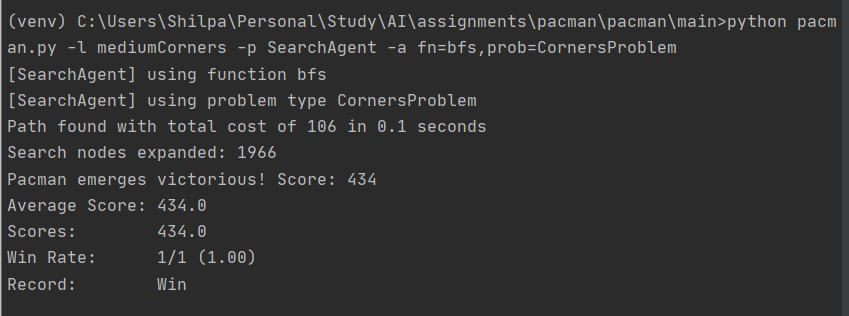
a. python pacman.py -l tinyCorners -p SearchAgent -a fn=bfs,prob=CornersProblem



**Observation:** As mentioned in the question, this algorithm takes a total cost of 28 for tinyCorners

1. **python pacman.py -l mediumCorners -p SearchAgent -a fn=bfs,prob=CornersProblem**





**Observation:** As mentioned in the question, this algorithm takes less than 2000 node expansion for the mediumCorners

**QUESTION 6.:**

In order for the Corner Heuristic to be consistent, we do the following analysis:

We know that if the calculated heuristic value of a node <= the calculated heuristic of its neighbour + cost of path from that neighbour to the node , then the heuristic is consistent. Also the proof of heuristic as per question is to show that UCS and Astar algorithms must return paths of same length.

The analysis based on this is given below:

* Here, heuristic used is the Manhattan distance to the goal.
* Table to show comparison:

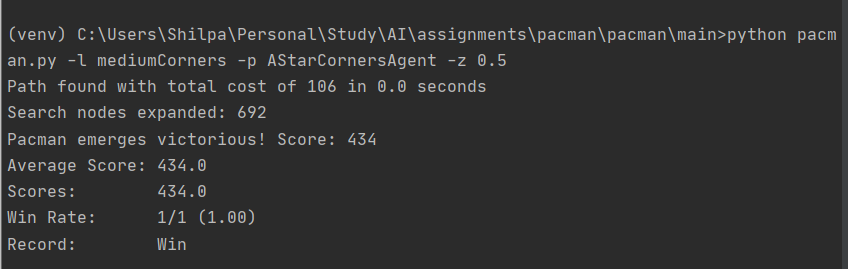
|  |  |  |
| --- | --- | --- |
| Algorithm | Cost | Nodes expanded |
| UCS | 106 | 1966 |
| A-star | 106 | 692 |

* Conclusion: Even though the total number of nodes expanded are different, the path to the goal state has same cost. Hence for the same problem, the heuristic algorithm using Manhattan distance is consistent.

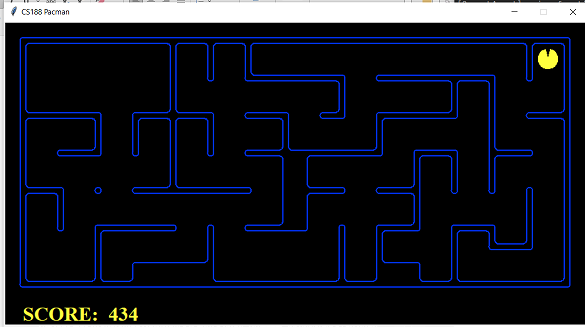
The execution commands and screenshots are shows below:

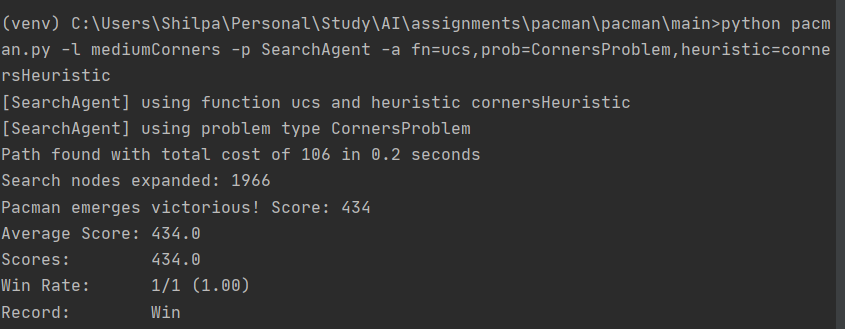
1. **python pacman.py -l mediumCorners -p AStarCornersAgent -z 0.5**





**UCS: python pacman.py -l mediumCorners -p SearchAgent -a fn=ucs,prob=CornersProblem,heuristic=cornersHeuristic**





**QUESTION 7:**

This algorithm has been tested on two Searches- Test Search and Tricky Search as per question and the analysis results and observations are discussed first. Later the screenshots of the run has also been shown :

**Comparison Table:**

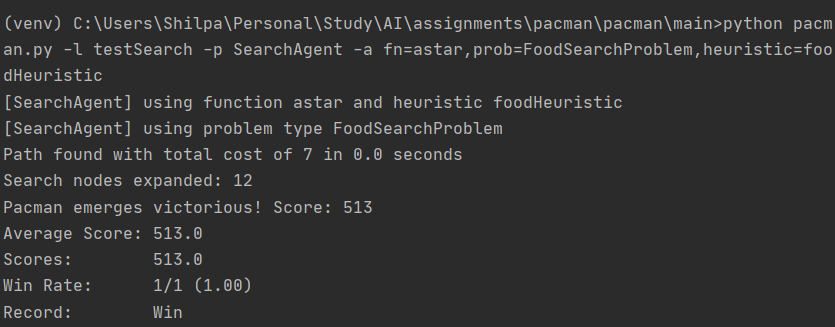
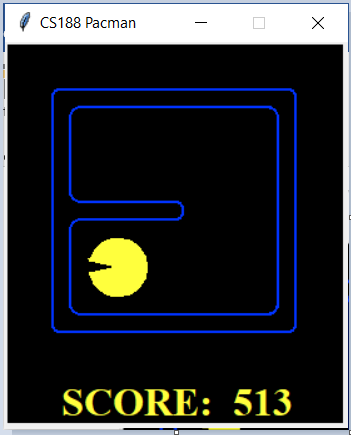
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Food Placement Type | Search Algorithm | Cost | Number of Nodes | Time Taken |
| testSearch | Astar | 7 | 12 | 0.0 seconds |
| testSearch | UCS | 7 | 14 | 0.0 seconds |
| trickySearch | Astar | 60 | 9551 | 6.3 seconds |
| trickySearch | UCS | 60 | 16688 | 21.7 seconds |

**Observations:**

1. We have found UCS slowing down as the problem got more complex even though the cost remains same, the number of nodes expanded and time taken is better in case of A-star than UCS.

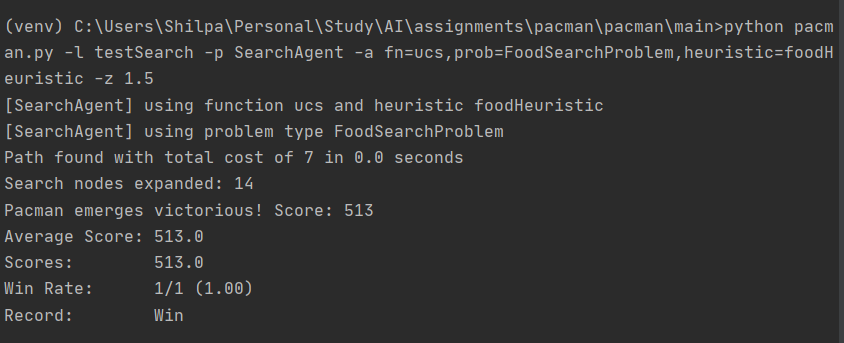
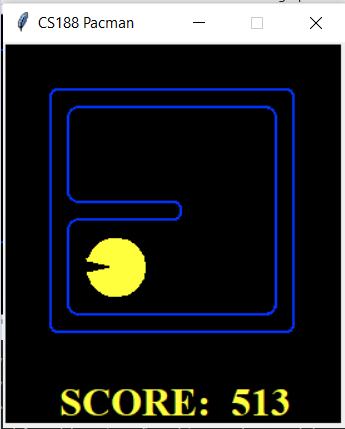
**TestSearch- Astar:**

1. **python pacman.py -l testSearch -p SearchAgent -a fn=astar,prob=FoodSearchProblem,heuristic=foodHeuristic**



1. **Test Search- UCS:**

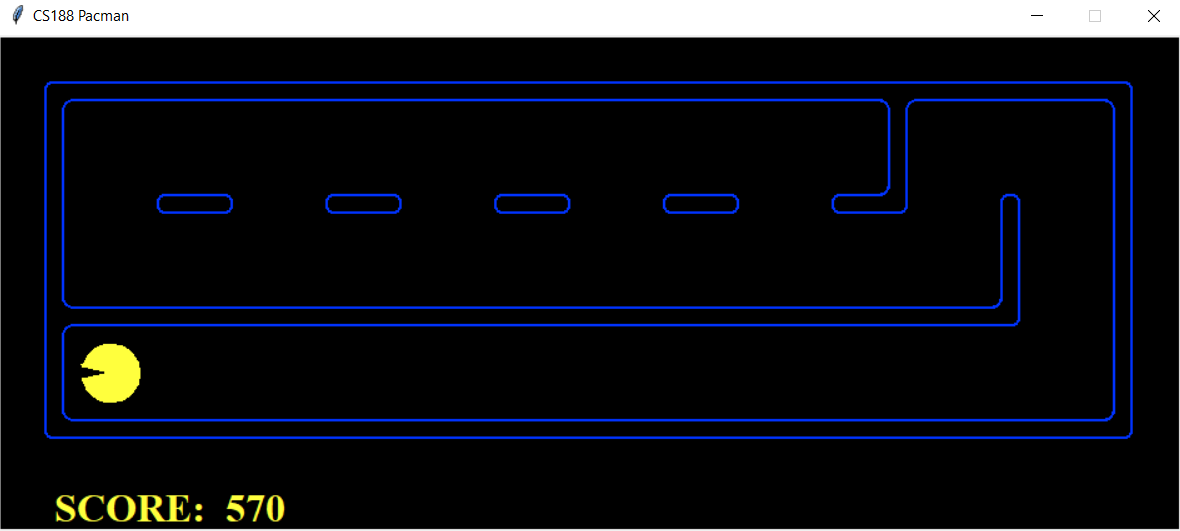
**python pacman.py -l testSearch -p SearchAgent -a fn=ucs,prob=FoodSearchProblem,heuristic=foodHeuristic -z 1.5**

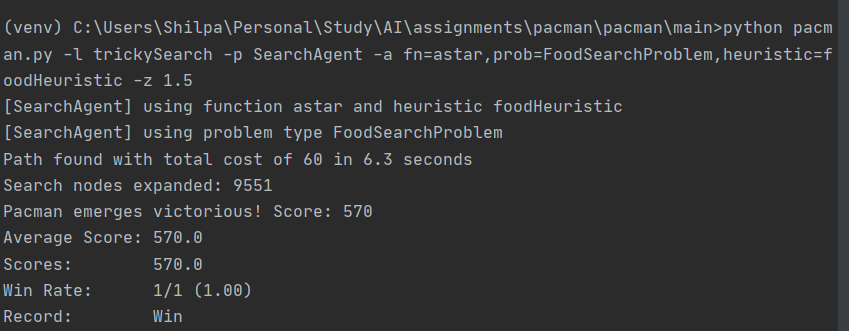


**Tricky Search:**

1. **Tricky Search Sstar:**

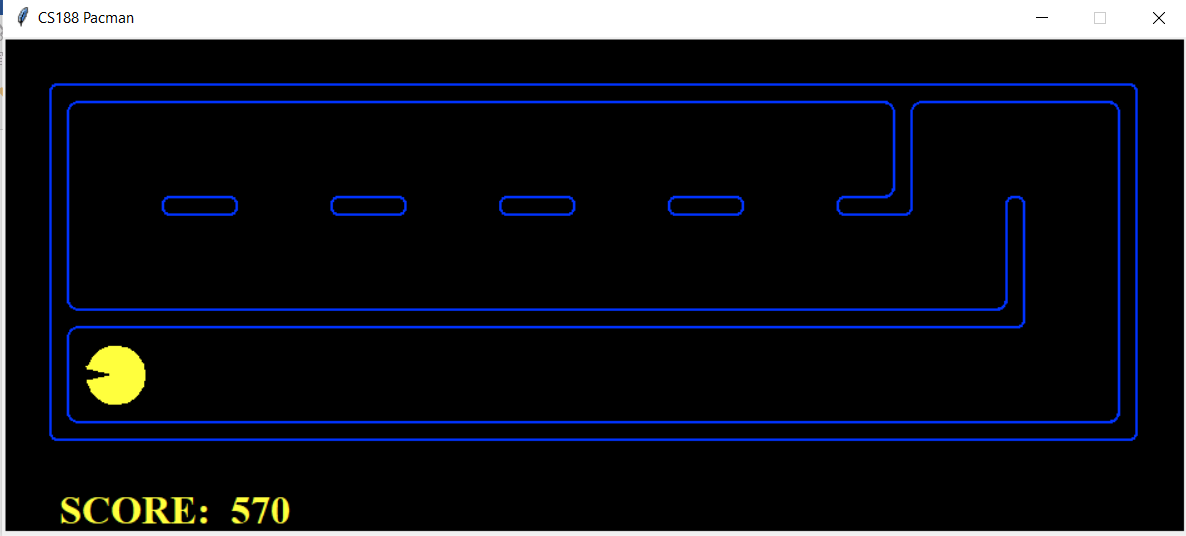
**python pacman.py -l trickySearch -p SearchAgent -a fn=astar,prob=FoodSearchProblem,heuristic=foodHeuristic -z 1.5**

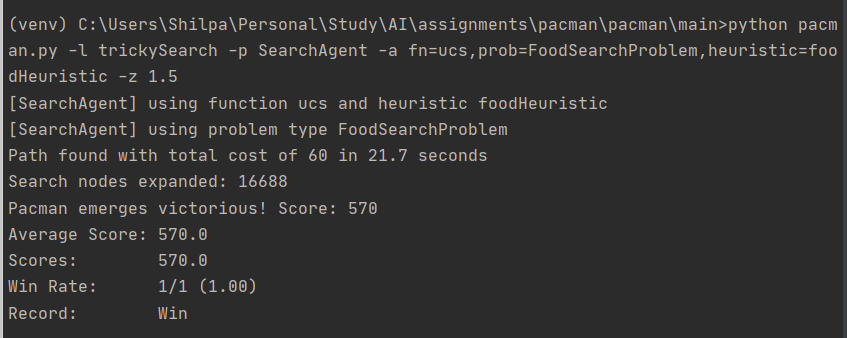




1. **Tricky Search- UCS:**

**python pacman.py -l trickySearch -p SearchAgent -a fn=ucs,prob=FoodSearchProblem,heuristic=foodHeuristic -z 1.5**

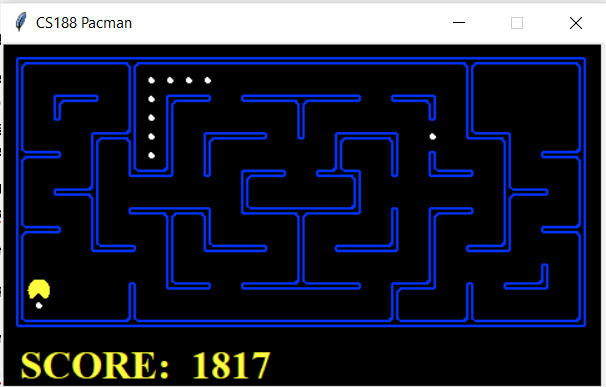
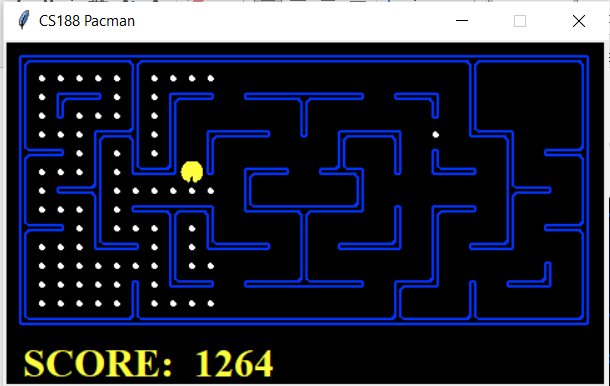




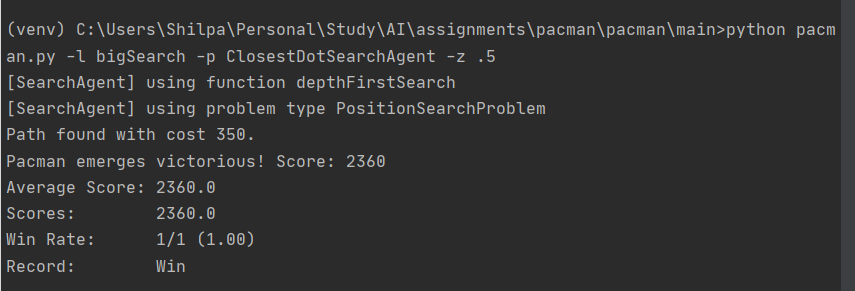
**QUESTION 8:**

In the given question, I have been asked to select a search algorithm which will make the Food search problem a greedy one in terms of selecting closest path to the food. That means we need to find a shortest route optimal algorithm in searching food dot so that the entire food search function becomes a greedy algorithm. Therefore, A-star should be our solution because A-star algorithm is optimal in nature and hence guaranteed to return the lowest cost path to a food. The function gets this return value and chooses the path accordingly in a greedy manner.

In answer to the next question as to why ClosestDotSearchAgent wont always find the shortest possible path through the maze. Let us try to understand from the below screenshots, taken at a time during running the command: python pacman.py -l bigSearch -p ClosestDotSearchAgent -z .5



In the second scenario (picture with score 1817) shown above, the agent has eaten all the closest point dots and reached the bottom left. However, there are two clusters of dot which stayed behind without visiting because none of these dots fell under the closet dot category when the agent was crossing that area (picture with score 1264). Sadly, now the agent has to go back a huge distance till the nearest bunch of points first (whose surroundings have already been covered once), process all the nodes in that bunch. Then again, the agent has to traverse to the second unexplored area containing a single dot. This will result in unnecessary addition to path cost of retraversal in the same area making the algorithm suboptimal.



As mentioned in the question, indeed the cost is 350.

**Acknowledgements:**

1. Artificial Intelligence: A Modern Approach, by Russell and Norvig, 4th edition.

2. <https://github.com/vishwapardeshi/AI-Pacman-Projects>

3. <https://github.com/karlapalem/UC-Berkeley-AI-Pacman-Project>