Foundations of Intelligent Systems

LAB 1

Rolling Die Mazes

Group 7:

Ashish Paralkar

Brij Shah

1 Index

2 Definition:

As described in the problem statement at https://www.cs.rit.edu/~ago/courses/630/lab1/index.html , the task is to implement A* search algorithm in a grid with blanks and walls with movements restricted by a die configuration (6 faces).

The constraints quoted from the above link:

1. The die begins with 1 on top, 2 facing up/north, and 3 facing right/east.
2. All opposite die faces add to 7 (for example: 1 on top + 6 on bottom = 7).
3. The number 6 should never be on top of the die facing 'up' (away from the grid).
4. The number 1 must be on top of the die when the goal location is reached.

So the task is to find shortest path (minimum cost in terms of distance) from Start position to goal position, with the above constraints enforced, using the A* algorithm which works like Djikstra's algorithm but makes use of an additional cost determining factor, or a 'heuristic'. We have to make use of 3 such admissible and consistent heuristics which will be described in the subsequent sections.

A state in the space of all states will be a place in the grid, defined by its x and y coordinates, the x and y coordinates of its parent, the die configuration (what digit its top, bottom, north, south, left and right faces show), and the direction the die moved in to reach the current spot(x,y). Note that top left corner is 0,0 and in a grid of size 4*4, the bottom right corner would be 3,3.

3 Admissible Heuristics

The first heuristic chosen, is the Euclidean distance of a spot from the goal spot. In a 2D space distance between 2 point x,y and a,b is

$$dist((x, y), (a, b)) = \sqrt{(x - a)^2 + (y - b)^2}$$

This is obviously an admissible heuristic because the lesser the Euclidean distance, the closer you are physically to your goal position. This is also a consistent heuristic because the diagonal distance is guaranteed to be less than the actual distance of reaching a goal position, dodging the walls in between in the grid.

The second Heuristic is the Manhattan distance. In a 2D space distance between 2 point x,y and a,b is

$$dist((x, y), (a, b)) = |(x - a)| + |(y - b)|$$

This is also an admissible heuristic because as with Euclidean distance, lesser the Manhattan distance, closer you are to your goal position, it is a monotonic function. It is also consistent because the again, the scalar addition of x and y distances is always less than or equal to the actual cost of reaching a goal in a grid that has walls which may cause you to divert from a path that you would otherwise choose in a completely free grid.

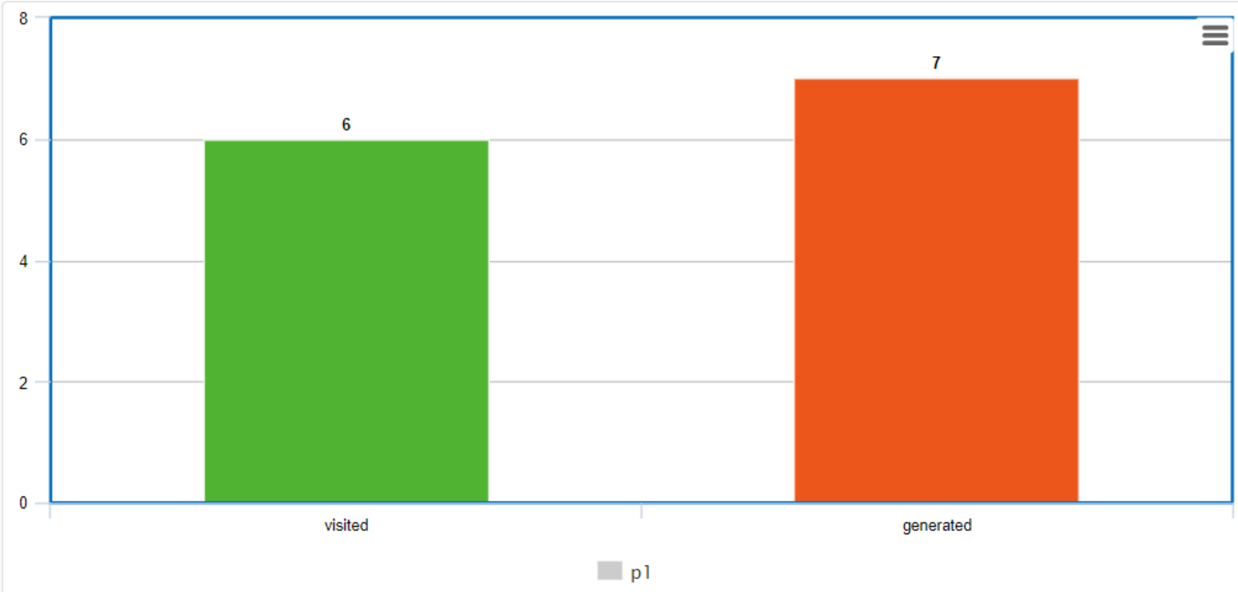
Having said this, we choose the 3rd heuristic as the average of the Manhattan and Euclidean distances. As mentioned above the consistency and admissibility holds because the following relation is true:

Actual calculated cost in a grid with walls >= Manhattan distance >= average of Manhattan and Euclidean distance >= Euclidean distance. (Its easily understood that equality holds in a special case : your start and goal positions are in a straight horizontal or vertical line and no walls in between.)

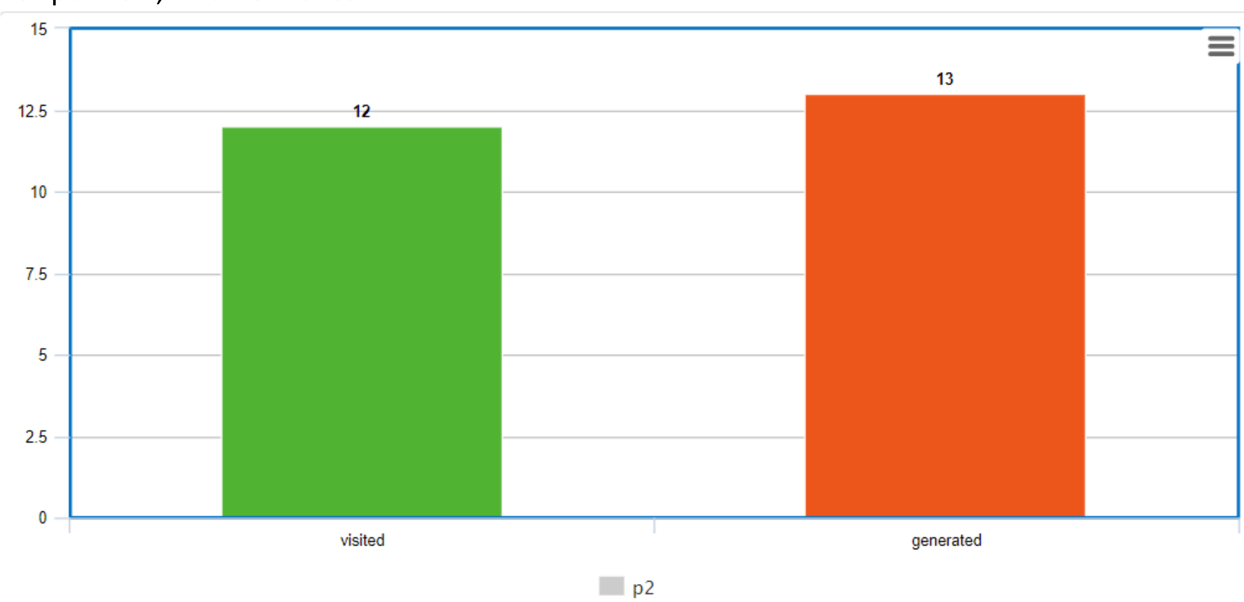
4 Performance metrics

Green indicates nodes visited, orange indicates nodes generated

For puzzle 1, all 3 heuristics:

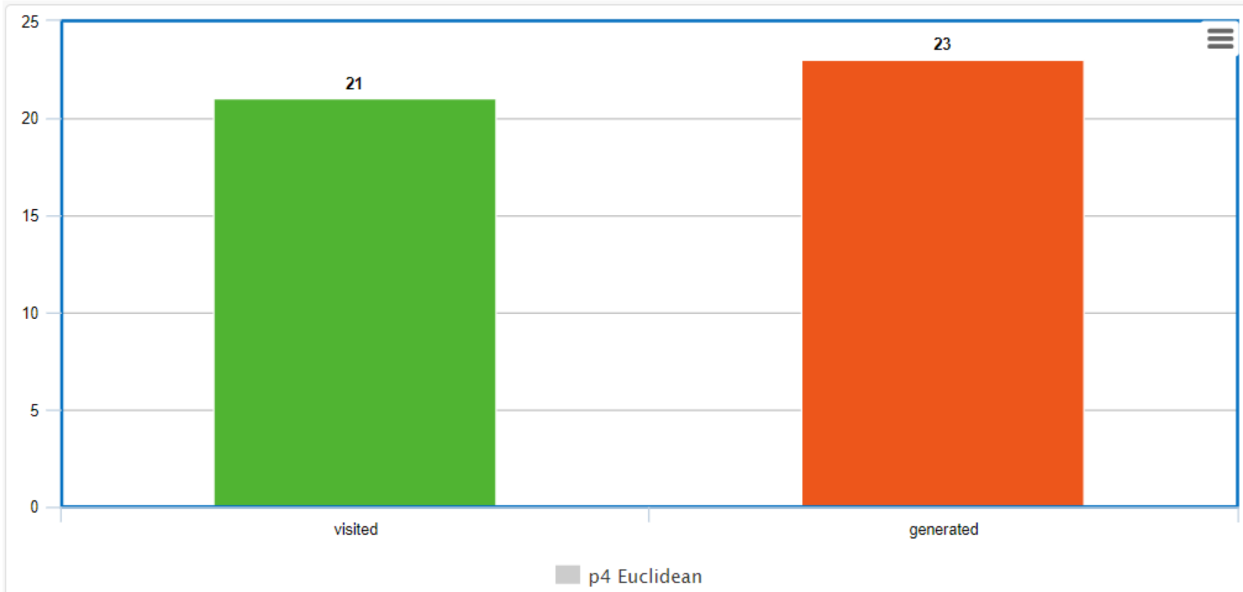


For puzzle 2, all 3 heuristics:

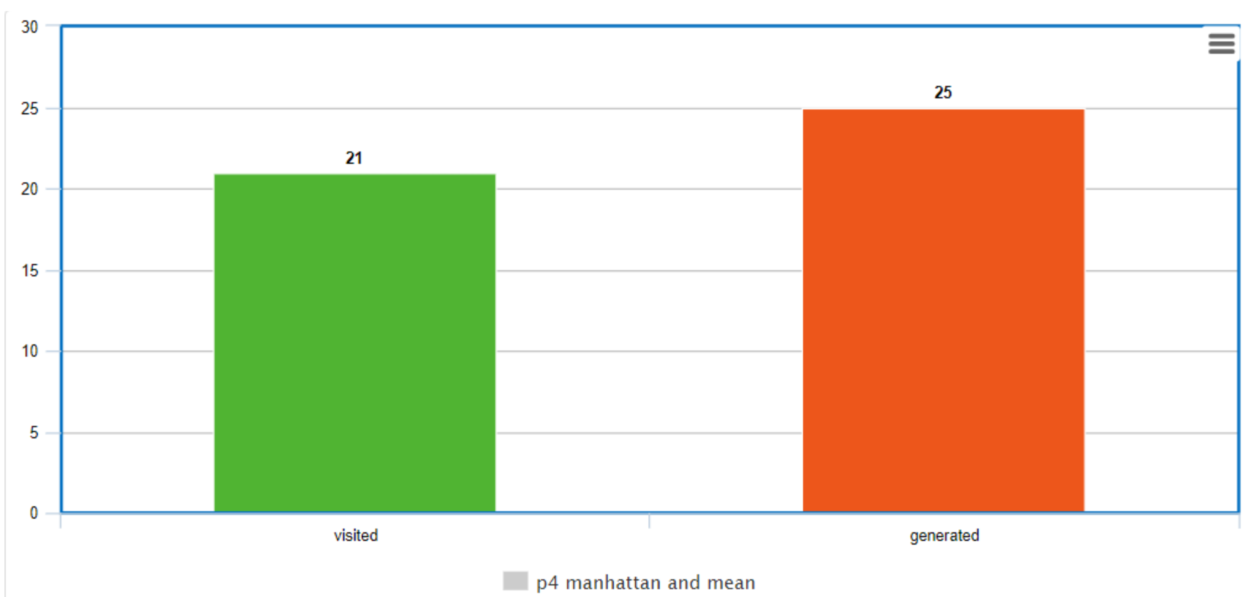


For puzzle 3 : no solution

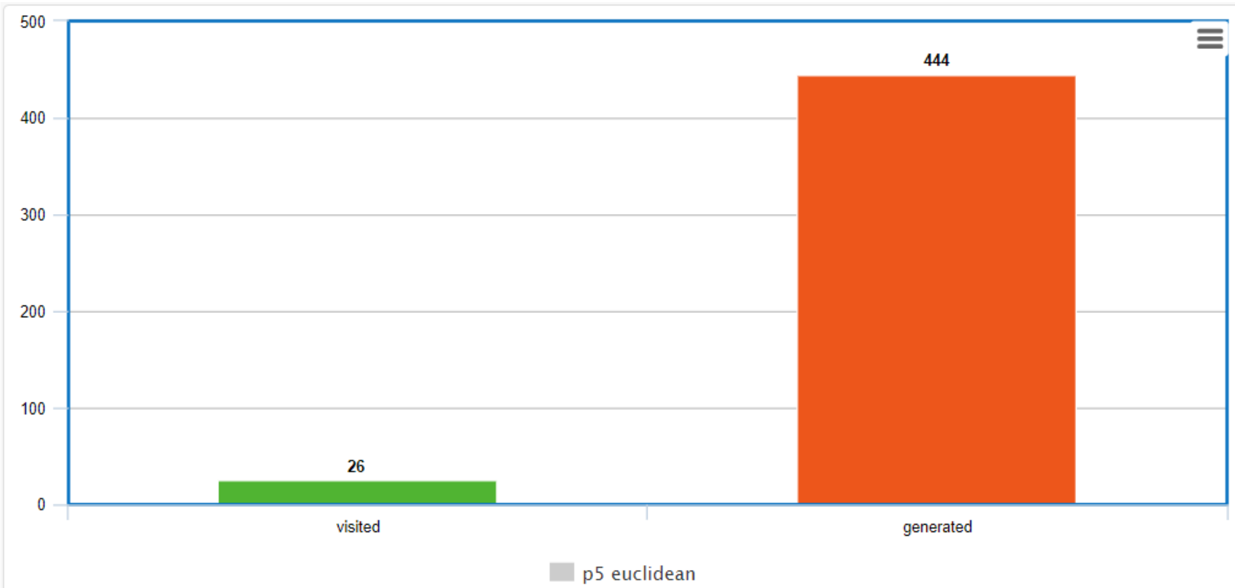
For puzzle 4 :
Euclidean Heuristic



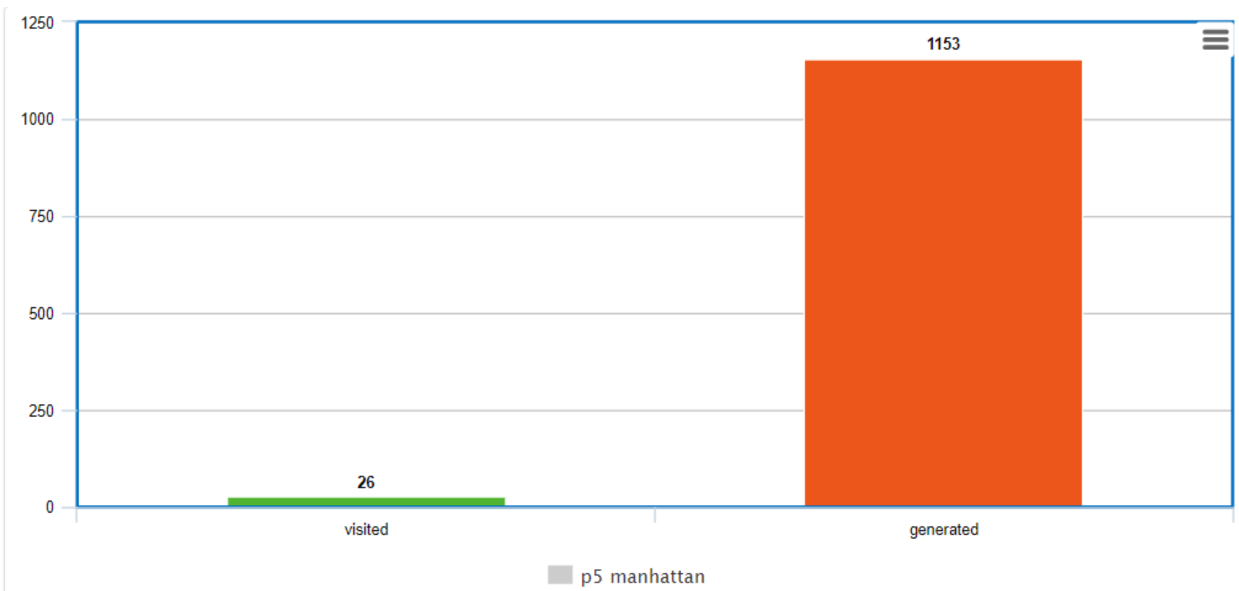
Manhattan and mean heuristic:



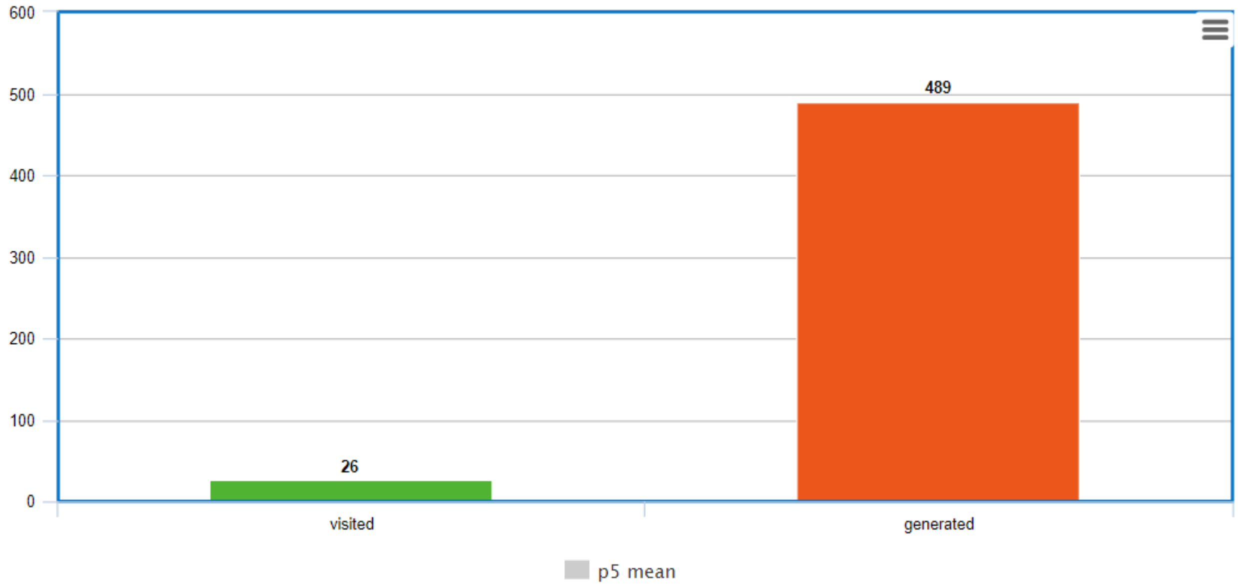
For puzzle 5:
Euclidean



Manhattan:



Mean heuristic:



5 Results

The results are pretty much as expected : The 3 chosen heuristics may have different number of nodes discovered as there is a definite variation between how Euclidean distance is defined, and how Manhattan distance is defined (and thus how their mean is defined) but they are admissible in the same regard (i.e. SOME function of (x distance, y distance)), and hence the number of nodes visited (and also the path generated) is Exactly the same for all 3.

Also one observation here seems to be that (going by puzzle 5's results especially) the Euclidean heuristic is more time efficient because it generates lesser discovered nodes than the mean heuristic and far lesser than the Manhattan heuristic.