CSE 4633 Programming Assignment 1

Naveen Neelakandan (nn149)

Part 1:

A* is a graph search algorithm which is widely used in AI. It is an extension of Dijkstra's algorithm and achieves better performance than Dijkstra's by using heuristics. Nodes are evaluated by combining g(n), the cost to a node and h(n), the estimated cost of the cheapest path to goal. The sum of g(n) and h(n) gives f(n) , the estimated cost of the cheapest solution through n. A* is optimal if h(n) is an admissible heuristic. That is, the heuristic is always an underestimate of the actual cost to goal. The simplest admissible heuristic is that of the tiles out of place heuristic. It is a simple count of the number of tiles not in their original goal positions. This is an underestimate as for each tile not in its goal position, we need at least one step to take it to the goal. The heuristic is not very informed though. A more informed heuristic is one which gives a more accurate measure of the cost to goal. The tiles out of place heuristic is admissible but not informed as the heuristic is quite trivial and not a good representation of the cost to goal. Because of this , when we implement A* using the Tiles out of place heuristic, we see that we have to generate and expand a huge number of nodes as our search is not very intelligent and thus we are unable to determine accurately the best path to goal at each step as many nodes have the same f(n) value. Hence, we have to expand a great number of nodes. The Manhattan distance heuristic is the sum of the distance from goal of each tile in the puzzle. This heuristic is admissible as a tile cannot reach the goal in fewer steps than its Manhattan distance. Using the Manhattan heuristic, we see a dramatic improvement in the number of nodes generated/ expanded. We now generate far fewer nodes as our search heuristic is more intelligent and thus can differentiate between states better to choose the shorter path to goal. The third heuristic is the Manhattan distance with linear conflict heuristic. It is an improvement to the Manhattan distance heuristic as we take into account the fact that 2 tiles already in their correct row/column must pass each other to get to the goal. We say that the Manhattan distance with linear conflict heuristic dominates the Manhattan distance heuristic as it is always greater for all nodes. We again see that the use of the more informed heuristic helps us to prune a greater deal of the search tree and so we generate/expand fewer nodes. Hence, from this experiment, we see how useful it is to have a more informed heuristic as we can prune a drastically greater amount of the search tree.