

# Solving 8-puzzle

## 1 Algorithm

### 1.1 A\* search

We implemented both programs using A\* search but with different heuristic functions. Given that our heuristic functions are consistent, the graph-search version of A\* search is known to be *complete*, *optimal* and *optimally efficient*. Therefore, it would be able to solve the 8-puzzle (if solvable).

## 2 Heuristics

We decided on the following heuristic functions:

- $h_1$ : Manhattan distance
- $h_2$ : Linear conflict

### 2.1 Manhattan distance

**Definition:**  $h_1$  is the sum of the distances of the tiles from their goal positions. Since diagonal moves are not allowed, the distance is the sum of horizontal and vertical distances.

For every node  $n$ , the step cost,  $c(n, n')$ , of getting to any of its successor node  $n'$  is 1 because for each move, only one tile is allowed to move one step closer to or further away from its goal position, i.e.  $h(n') = h(n) \pm 1$ . When we substitute step cost as 1 into  $h(n) \leq c(n, n') + h(n')$ , we get  $h(n) \leq 1 + h(n')$ . It is then trivial to see that the condition holds, therefore  $h_1$  is consistent.  $\square$

### 2.2 Linear conflict

**Definition:** Two tiles  $t_j$  and  $t_k$  are in a linear conflict if  $t_j$  and  $t_k$  are in the same line, the goal positions of  $t_j$  and  $t_k$  are both in that line,  $t_j$  is to the right of  $t_k$ , and the goal position of  $t_j$  is to the left of the goal position of  $t_k$ .

To derive the Linear Conflict estimate for any node  $n$ ,

1. Calculate the minimum number of tiles that must be removed from  $row_1$  such that there are no more linear conflicts.
2. Repeat Step 1 for the other rows and columns and sum them.
3.  $LinearConflict(n) = 2 * \text{result from Step 2}$

Using the above property, we augment the manhattan distance heuristic by adding the linear conflict estimate to it, thereby making it a more informed heuristic.

$$h_2(n) = ManhattanDistance(n) + LinearConflict(n) = h_1(n) + LinearConflict(n)$$

*Proof of consistency for  $h_2$ .*

Let  $md(n, x)$  be the manhattan distance of tile  $x$  in node  $n$  and  $lc(n, r_i)$  be the number of tiles that must be removed from row  $r_i$  to resolve linear conflicts. Assume that tile  $x$  is moving from row  $r_{old}$  to  $r_{new}$ , while remaining in column  $c_{old}$ . The impact of the change in position of tile  $x$  relative to its goal position can be split into three scenarios:

1. The goal position of  $x$  is in neither row.  $md(n', x) = md(n, x) \pm 1$ . There are no new linear conflicts. Hence,  $h_2(n') = h_2(n) \pm 1$  and  $h_2(n') + c(n, n') = h_2(n') + 1 \geq h_2(n)$ .

2. The goal position of  $x$  is in  $r_{new}$ . Since  $x$  moved into the row containing its goal position,  $md(n', x) = md(n, x) - 1$  and this may or may not have new linear conflicts in row  $r_{new}$ , so  $lc(n', r_{new}) = lc(n, r_{new})$  or  $lc(n', r_{new}) = lc(n, r_{new}) + 2$ . Because  $r_{old}$  is not the goal row of  $x$ , its presence would not have contributed to any linear conflict there so  $lc(n', r_{old}) = lc(n, r_{old})$ . Hence,  $h_2(n') = h_2(n) \pm 1$  and  $h_2(n') + c(n, n') = h_2(n') + 1 \geq h_2(n)$ .
3. The goal position of  $x$  is in  $r_{old}$ . Since  $x$  moved out of the row containing its goal position,  $md(n', x) = md(n, x) + 1$  and we do not know whether it originally contributed to linear conflicts in  $r_{old}$  so  $lc(n', r_{old}) = lc(n, r_{old})$  or  $lc(n', r_{old}) = lc(n, r_{old}) - 2$ . Because  $r_{new}$  is not the goal row of  $x$ , its presence would not have contributed to any linear conflict there so  $lc(n', r_{new}) = lc(n, r_{new})$ . Hence,  $h_2(n') = h_2(n) \pm 1$  and  $h_2(n') + c(n, n') = h_2(n') + 1 \geq h_2(n)$ .

In all three scenarios,  $h_2$  remains consistent. By symmetry of the 8-puzzle,  $h_2$  is similarly consistent for movements from column to column.  $\square$

The above proof is adapted from a paper published by Hansson, Mayer and Yung in 1985 on *Generating Admissible Heuristics by Criticizing Solutions to Relaxed Models*.

### 3 Statistics

For given inputs, report number of nodes generated, maximum size of frontier that is reached. Check that solution is within 300 moves.

### 4 Analysis