HEURISTICS

① So how do we come up with heuristics? In practice, it tends to be difficult to think about consistency but relatively easy to think about admissibility, so we come up with admissible heuristics and then check that (or hope that) they're consistent.

② Let's consider some examples. In word ladder, the heuristic we've been using is the number of letters that are different from the target word.

e.g. TRAIN

PRAWN

It's fairly clear that this is a lower bound on the cost of the optimal completion path, because we have to AT LEAST change every incorrect letter to reach the target word.

HEURISTICS

③ A good example problem that demonstrates the engineering of heuristics is the 8-puzzle.

The object of an 8-puzzle is to slide tiles until you match a goal configuration, e.g.

7	2	4
5		6
8	3	1

	1	2
3	4	5
6	7	8

A natural state machine formulation defines its states as puzzle configurations, and its actions as L, R, U, D (move a tile L/R/U/D into the empty square)

4	6	2
8	1	3
7	5	

L

R

4	6	2
8	1	3
7		5

D U

4	6	2
8	1	
7	5	3

HEURISTICS

(4) One admissible heuristic is the number of misplaced tiles:

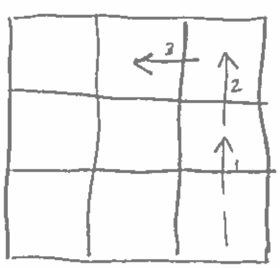
7	2	4
5		6
8	3	1

All 8 are misplaced, so $H\left(\begin{array}{|c|c|c|}\hline 7 & 2 & 4 \\ \hline 5 & & 6 \\ \hline 8 & 3 & 1 \\ \hline\end{array}\right) = 8.$

This has to be a lower bound on H^*, because we can only move one tile at a time, and each misplaced tile has to move at least once.

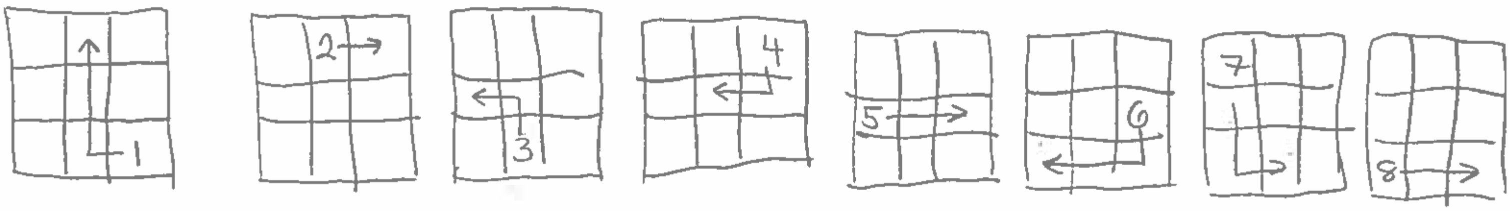
---

(5) Following this logic, we know that we need to move each misplaced tile to its correct position at some point, which will take at least k moves, where k is the "Manhattan distance" from the tile's current position to its goal position



Manhattan distance
$= 3$

⑥ So another admissible heuristic is the sum of the Manhattan distances of each tile to its goal position:



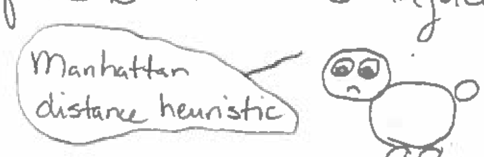
$H(q) = 3 + 1 + 2 + 2 + 2 + 3 + 3 + 2$
$= 18$

⑦ A good strategy for creating admissible heuristics is to write down the actions of the state machine we want to solve. For instance, 8-puzzle can be described:

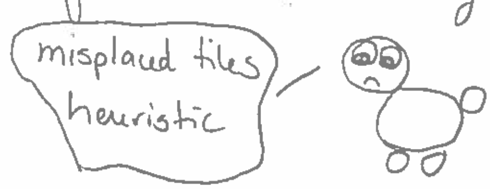
A tile can move from square A to square B if
A is adjacent to B and B is blank.

Then we can make the problem easier by removing constraints.

A tile can move from square A to square B if A is adjacent to B ~~and B is blank~~



A tile can move from square A to square B ~~if A is adjacent to B and B is blank~~.



Heuristics

(8) The key is, a solution to an easier version of our original problem is an admissible heuristic for the original problem.

Consider again the word ladder problem's basic action:

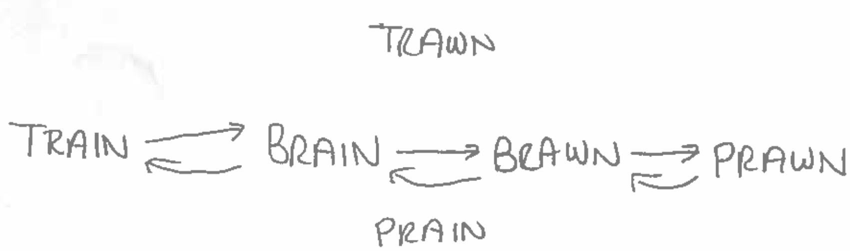
We can replace a letter if the resulting word is in the dictionary.

We can instead solve the simpler problem, where:

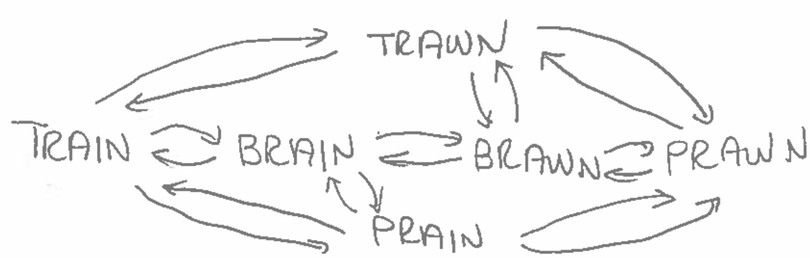
We can replace a letter ~~if the resulting word is in the dictionary~~

This gives us the "number of incorrect letters" heuristic.

(9) Another way to view this is that we're creating a "supergraph" of our state machine and solving that to create our heuristic:



becomes



HEURISTICS

(10) Let's revisit the description of the 8-puzzle's action:

A tile can move from square A to square B if
A is adjacent to B and B is blank.

There's a third heuristic we can derive from this:

A tile can move from square A to square B if
~~A is adjacent to B and~~ B is blank.

(11) When deriving a new heuristic, the first question should ask is: is it efficiently computable? If it's just as hard to compute the relaxed problem as the original, then there's no point.

Fortunately, we can solve this efficiently. Suppose we're in state:

	1	2
3	4	5
6	8	7

If the "blank" is in the correct position, then move any misplaced tile into its position:

8	1	2
3	4	5
6		7

Otherwise, move the correct tile to the blank's position:

8	1	2
3	4	5
6	7	

, then

	1	2
3	4	5
6	7	8

HEURISTICS

(12) The second question to ask is: is this any better than other heuristics I already know? Well...

	1	2
3	4	5
6	8	7

misplaced tiles heuristic = 2

Manhattan dist heuristic = 2

jump-to-blank heuristic = 3

So yes, it can be better, sometimes considerably better:

	2	1
4	3	8
7	6	5

misplaced tiles heuristic = 8

Manhattan dist heuristic = 8

jump-to-blank heuristic = 12

(13) The final question to ask is: is this ALWAYS better than other heuristics I already know?

— if so, we say the heuristic <u>dominates</u> the other heuristics

Not in this case:

8	1	2
3	4	5
6	7	

Manhattan dist heuristic = 4

jump-to-blank heuristic = 1

HEURISTICS

(14) If we have two admissible heuristics H_1 and H_2, we can make a stronger heuristic by combining them:

$$H(q) = \max\{H_1(q), H_2(q)\}$$

For any state q, if $H(q) = H_1(q)$, then $H(q) = H_1(q) \leq H^*(q)$
if $H(q) = H_2(q)$, then $H(q) = H_2(q) \leq H^*(q)$

So the new heuristic H is also admissible.