Informed Search 2

Admissible Heuristic, Proof of Optimality of A*, Properties of Heuristic Function

Admissible heuristic

- A heuristic function h(n) is admissible if for every node n, h(n)<=h*(n), where h*(n) is the true cost to reach the goal state from n.
- An admissible heuristic never overestimates the actual cost to reach the goal. Admissible heuristic is optimistic by nature.
- If h(n) is admissible, A* using tree search is optimal.

Monotonicity

- A heuristic is consistent if: h(n) <= cost(n,n’)+h(n’) , where n’ is the neighbor of n.
- If a heuristic h is consistent the f values along any path will be non decreasing:

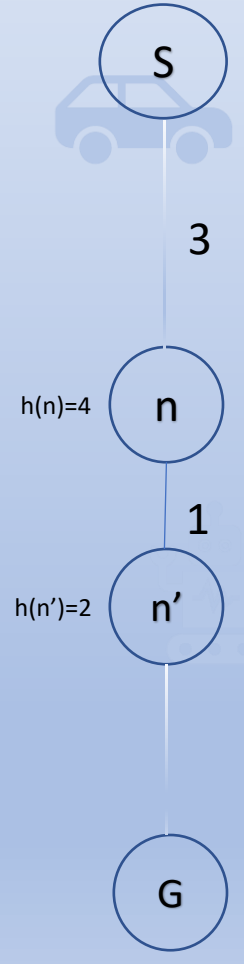
f(n')=estimated distance from start to goal through n’

=g(n)+cost(n,n')+h(n’)

>=g(n)+h(n)

=f(n)

So f never decreases along the path



Inconsistent

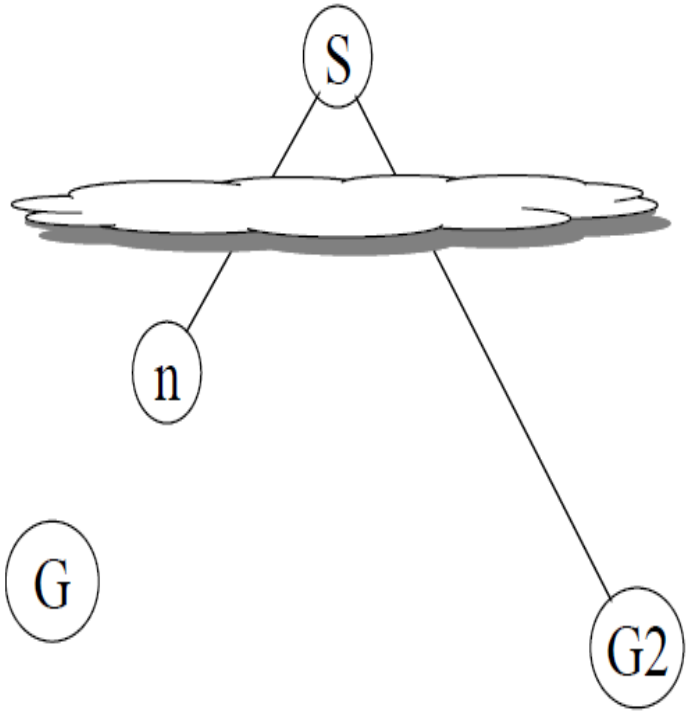
Monotonicity

- A monotone heuristic is such that along any path the f-cost never decreases.
- But if this property does not hold for a given heuristic function, we can make the f value monotone by making use of the following trick (m is a child of n)

$$f(m) = \max (f(n), g(m) + h(m))$$

Proof of Admissibility of A*

- Let G be an optimal goal state
- C* is the optimal path cost.
- G2 is a suboptimal goal state: g(G2) > C*
- Suppose A* has selected G2 from OPEN for expansion.
- Consider a node n on OPEN on an optimal path to G. Thus C* ≥ f(n)
- Since n is not chosen for expansion over G2, f(n) ≥ f(G2)
- G2 is a goal state. f(G2) = g(G2)
- Hence C* ≥ g(G2).
- This is a contradiction. Thus A* could not have selected G2 for expansion before reaching the goal by an optimal path.



Proof of Completeness of A*

- Let G be an optimal goal state.
- A* cannot reach a goal state only if there are infinitely many nodes where $f(n) \leq C^*$.
- This can only happen if either happens:
  - There is a node with infinite branching factor. The first condition takes care of this.
  - There is a path with finite cost but infinitely many nodes. But we assumed that Every arc in the graph has a cost greater than some $\varepsilon > 0$. Thus if there are infinitely many nodes on a path $g(n) > f^*$, the cost of that path will be infinite.

Lemma: A* expands nodes in increasing order of their f values.

- A* is thus complete and optimal, assuming an admissible and consistent heuristic function (or using the pathmax equation to simulate consistency).
- A* is also optimally efficient, meaning that it expands only the minimal number of nodes needed to ensure optimality and completeness.

Performance Analysis of A*

- In practice most heuristics have proportional error.
- It becomes often difficult to use A* as the OPEN queue grows very large.
- A solution is to use algorithms that work with less memory.

Properties of Heuristics

Dominance:

- h2 is said to dominate h1 iff h2(n) ≥ h1(n) for any node n.
- A* will expand fewer nodes on average using h2 than h1.

Proof:

- Every node for which f(n) < C* will be expanded. Thus n is expanded whenever h(n) < f* - g(n)
- Since h2(n) ≥ h1(n) any node expanded using h2 will be expanded using h1.

Properties of Heuristics

- Suppose you have identified a number of non-overestimating heuristics for a problem: h1(n), h2(n), … , hk(n)
- Then max (h1(n), h2(n), … , hk(n))
- is a more powerful non-overestimating heuristic. This follows from the property of dominance

Heuristic Function of 8-puzzle Problem

- H1=the numbers of tiles that are in the wrong position. It is admissible because any tiles that is out of place must be moved at least once.
- H2= the sum of the distances of the tiles from their goal positions. The distance will be count as the sum of the horizontal and vertical distances. This is called Manhattan distance.
- H2 is admissible because any move can only move one tile one step closer to the goal.