CS3243 Tutorial Group 7

Tutorial 3 – Informed Search

Quick recap

- Class of algorithms: best-first search
  - Use an evaluation function f(n) = cost estimate; expand node with the lowest estimate first

- Greedy best-first search
  - Evaluation function f(n) = h(n)
  - Heuristic function h(n) = estimated cost of cheapest path from n to goal
  - Complete for graph-search (in finite spaces) but **not complete for tree-search** (even in finite spaces)

- A* search: avoid expanding paths that are already expensive
  - Evaluation function f(n) = g(n) + h(n)

- Heuristics:
    - Admissible (never overestimates)
    - Consistent/monotone = f(n) is monotonically non-decreasing along every path
    - Dominance: dominant heuristics expand fewer nodes
    - Deriving admissible heuristics from relaxed problems

- Local search: hill-climbing

Intuition for admissible vs consistent

- **admissible** if it always underestimates path costs to the goal:

$$\forall n,$$
$$h(n) \leq h^*(n)$$

- **consistent** if it always underestimates step costs between nodes:

$$\forall n, n',$$
$$h(n) \leq d(n, n') + h(n')$$

(original statement)

Equivalently:

$$h(n) - h(n') \leq d(n, n')$$

(restatement)

implied estimate of the step cost between n and n'

Now obvious why consistent implies admissible!

Simple proof that UCS is optimal (step costs $\geq \varepsilon$)

A* with consistent h(n) has the same proof

- We use these invariants of UCS, which follow directly from the construction of the algorithm:
  1. The frontier expands by exploring all edges out of the initial node, i.e. it forms a perimeter around it.
  2. Since all step costs $\geq \varepsilon$, $g(n)$ is **monotonically non-decreasing along any path**.
  3. When a node n is chosen for expansion, $g(n)$ is $\leq$ the cost of any other nodes in the frontier.

Consistency guarantees this; just replace g(n) with f(n) throughout

- Together, these invariants imply the property P:

whenever a node n is chosen for expansion, g(n) is the shortest path cost from the initial node to n

- Proof (by contradiction):
  - n is chosen for expansion.
  - Suppose there exists a path from the initial node to n that has lower cost than $\mathrm{g}(n)$.
  - By inv. 1, the only way to find such a path is by expanding one of the nodes in the frontier (expanding the perimeter).
  - But by inv. 3, all nodes q in the frontier have $g(q) \leq g(n)$, since n was selected for expansion.
  - However, by inv 2., all paths from any node q have cost at least $\mathrm{g}(q) + \varepsilon$, which is strictly greater than $g(n)$. Contradiction.
- By P, when a goal node is expanded, it has the lowest path cost
  - i.e. UCS is optimal