CS2109S Tutorial 2

Informed Search

(AY 24/25 Semester 2)

February 7, 2025

(Prepared by Benson)

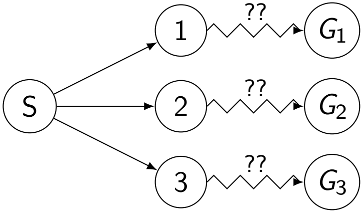
Contents

Informed Search

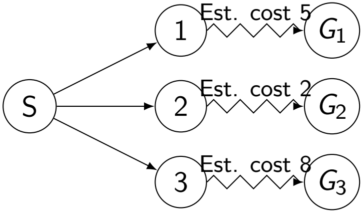
- Recap
- Q1. Pacman
- Q2. Euclidean Route-Finding
- Q3. Consistent $\Rightarrow$ Admissible
- Q4. Fagaras to Craiova
- Q5. Inconsistent Heuristic
- Bonus. A* Search vs Dijkstra

Recap: Informed Search

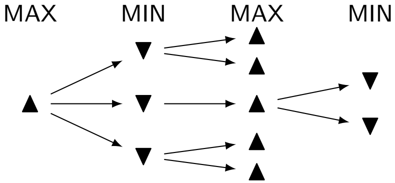
Uninformed search:



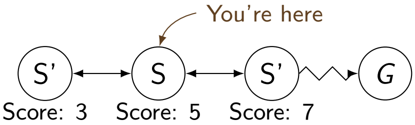
Informed search:



Adversarial search:

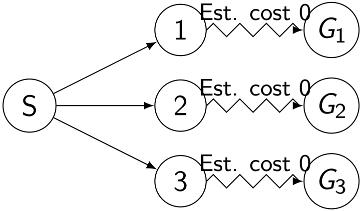


Local search:

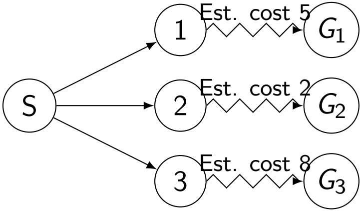


Recap: Informed Search

Searching with a **heuristic** (*estimates* the cost to the goal).



Degrades to uninformed search



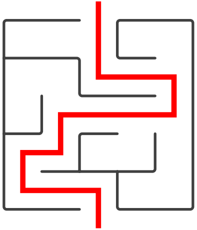
A good one?

Qn: Why would we prefer informed search over uninformed search?

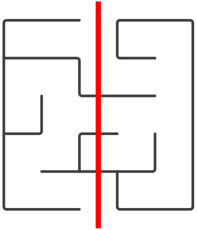
Recap: Informed Search

Admissible
$h(v) \leq h^*(v)$

$\Rightarrow$ Never overestimates the true cost of the **path**



True cost



Admissible heuristic
(**Relaxed problem**: We can go through walls.)

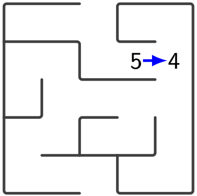
Recap: Informed Search

Consistent

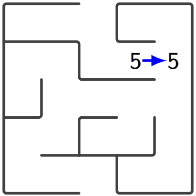
$$h(v) \leq c(v, a, u) + h(u)$$

i.e. $h(v) - h(u) \leq c(v, a, u)$

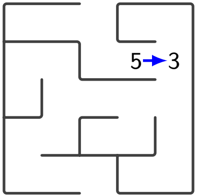
$\Rightarrow$ Never overestimates the cost of an **action**



OK



OK



Inconsistent

Recap: Informed Search

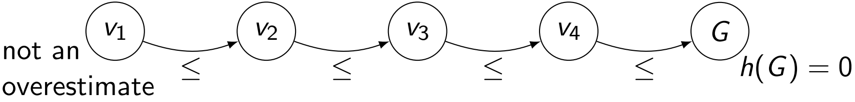
Admissible
$h(v) \leq h^*(v)$

$\Rightarrow$ Never overestimate cost of path

Consistent
$h(v) \leq c(v, a, u) + h(u)$

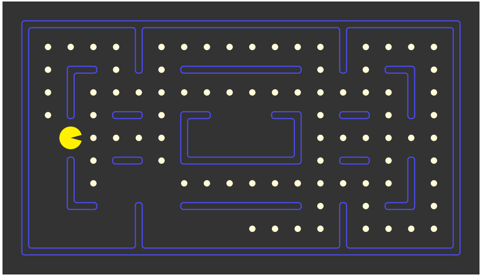
$\Rightarrow$ Never overestimate cost of action

Intuitive: Consistent $\Rightarrow$ Admissible



Q1. Pacman

Pacman has to eat all the pellets in the maze while executing the least amount of (4-directional) moves. Devise a non-trivial admissible heuristic for this problem.

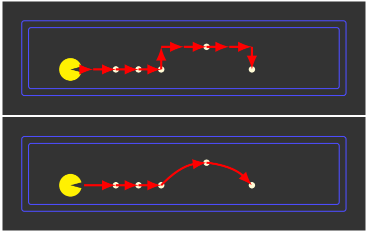


Q1. Pacman

- $h_1(n) =$ The number of pellets left.

Admissible

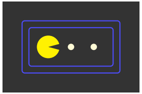
$$h(v) \leq h^*(v)$$



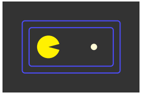
Relaxed problem: Pacman can jump to any cell in one move.

Consistent

$$h(v) \leq c(v, a, u) + h(u)$$



$\Delta h(n) = -1$



$\Delta h(n) = 0$

In both cases, the cost of the action (1) is not overestimated.

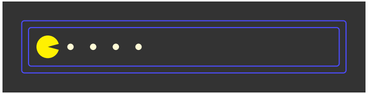
Q1. Pacman

- $h_2(n)$ = The **sum** of distances to all pellets.

Admissible

$h(v) \leq h^*(v)$

Counterexample:



- $h(v) = 1 + 2 + 3 + 4 = 10$
- $h^*(v) = 4$

Consistent

$h(v) \leq c(v, a, u) + h(u)$

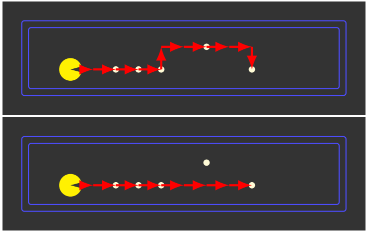
Since Consistent $\Rightarrow$ Admissible, we have Not Admissible $\Rightarrow$ Not Consistent.

CS1231/S Refresher: $X \Rightarrow Y$ is equivalent to $\sim X \Rightarrow \sim Y$ (contrapositive).

Q1. Pacman

- $h_3(n) =$ The **maximum** distance to any pellet.

Admissible
$h(v) \leq h^*(v)$



Relaxed problem: Pacman only needs to eat the farthest pellet.

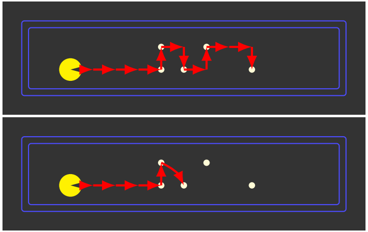
Consistent
$h(v) \leq c(v, a, u) + h(u)$

The maximum distance decreases by at most 1 after a move.

Q1. Pacman

$h_4(n)$ = The distance to the closest pellet + number of pellets adjacent to that pellet. (If there are multiple closest pellets, take one with the most adjacent pellets.)

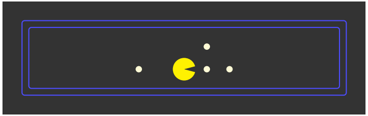
Admissible
$h(v) \leq h^*(v)$



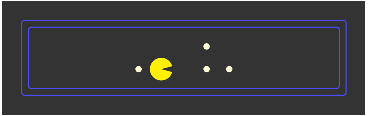
Relaxed problem: Pacman only needs to eat any pellet and its neighbours (by "jumping").

Consistent
$h(v) \leq c(v, a, u) + h(u)$

Counterexample:



- $h(v) = 1 + 2 = 3$



- $h(u) = 1 + 0 = 1$

Q2. Euclidean Route-Finding

Given a graph $G = (V, E)$ where each node v_n having coordinates (x_n, y_n), each edge (v_i, v_j) having weight equals to the distance between v_i and v_j, and a unique goal node v_g.

$$h_{SLD}(n) = \sqrt{(x_n - x_g)^2 + (y_n - y_g)^2}$$

$$h_1(n) = \max\{|x_n - x_g|, |y_n - y_g|\}$$

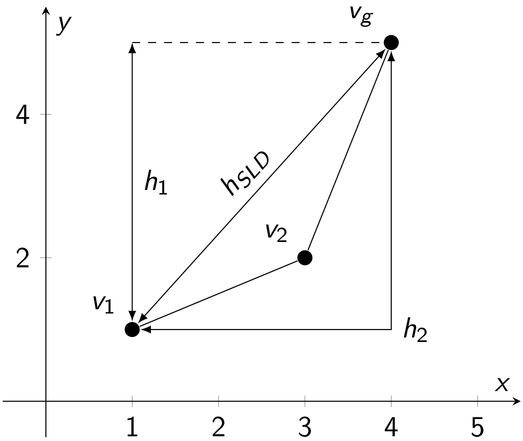
$$h_2(n) = |x_n - x_g| + |y_n - y_g|$$

(a) Is $h_1(n)$ an admissible and consistent heuristic?

(b) Is $h_2(n)$ an admissible heuristic?

(c) Which heuristic function would you choose for A* search?

Q2. Euclidean Route-Finding



- $h^*(v_1) = \sqrt{5} + \sqrt{10} = 5.39$
- $h_{SLD}(v_1) = \sqrt{3^2 + 4^2} = 5$
- $h_1(v_1) = \max\{3, 4\} = 4$
- $h_2(v_1) = 3 + 4 = 7$

Q2. Euclidean Route-Finding

$$h_{SLD}(n) = \sqrt{(x_n - x_g)^2 + (y_n - y_g)^2}$$
$$h_1(n) = \max\{|x_n - x_g|, |y_n - y_g|\}$$
$$h_2(n) = |x_n - x_g| + |y_n - y_g|$$

- Want to show: $h_1(n) \leq h_{SLD}(n)$ and hence h_1 is admissible.
- We have

$$\sqrt{(x_n - x_g)^2 + (y_n - y_g)^2} \geq \sqrt{(x_n - x_g)^2} = |x_n - x_g|$$

  and similarly

$$\sqrt{(x_n - x_g)^2 + (y_n - y_g)^2} \geq \sqrt{(y_n - y_g)^2} = |y_n - y_g|$$

  Therefore,

$$\max\{|x_n - x_g|, |y_n - y_g|\} \leq \sqrt{(x_n - x_g)^2 + (y_n - y_g)^2} \leq h^*(n)$$

Q2. Euclidean Route-Finding

$$h_{SLD}(n) = \sqrt{(x_n - x_g)^2 + (y_n - y_g)^2}$$

$$h_1(n) = \max\{|x_n - x_g|, |y_n - y_g|\}$$

$$h_2(n) = |x_n - x_g| + |y_n - y_g|$$

- Want to show: h_1 is consistent, i.e. $h_1(n) - h_1(n') \leq c(n, a, n')$.

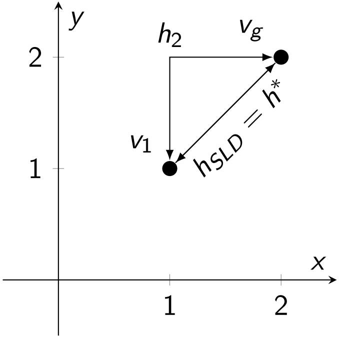
$$\begin{aligned} h_1(n) - h_1(n') &= \max\{|x_n - x_g|, |y_n - y_g|\} - \max\{|x_{n'} - x_g|, |y_{n'} - y_g|\} \\ &\leq \max\{|x_n - x_g| - |x_{n'} - x_g|, |y_n - y_g| - |y_{n'} - y_g|\} \\ &\leq \max\{|x_n - x_{n'}|, |y_n - y_{n'}|\} \\ &\leq \sqrt{(x_n - x_{n'})^2 + (y_n - y_{n'})^2} \blacktriangleleft \text{ by the previous slide} \end{aligned}$$

Q2. Euclidean Route-Finding

$$h_{SLD}(n) = \sqrt{(x_n - x_g)^2 + (y_n - y_g)^2}$$

$$h_1(n) = \max\{|x_n - x_g|, |y_n - y_g|\}$$

$$h_2(n) = |x_n - x_g| + |y_n - y_g|$$



- $h^*(v_1) = \sqrt{2} = 1.41$
- $h_2(v_1) = 1 + 1 = 2$
- $\therefore$ h_2 overestimates the cost to the goal, not admissible.

Q2. Euclidean Route-Finding

$$h_{SLD}(n) = \sqrt{(x_n - x_g)^2 + (y_n - y_g)^2}$$

$$h_1(n) = \max\{|x_n - x_g|, |y_n - y_g|\}$$

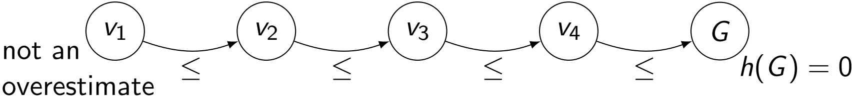
$$h_2(n) = |x_n - x_g| + |y_n - y_g|$$

Which heuristic function would you choose for A* search?

- h_2 is not admissible $\Rightarrow$ cannot be used.
- h_{SLD} dominates h_1 $\Rightarrow$ better choice.
  - $h_{SLD}(n) \geq h_1(n)$ for all states n.

Q3. Consistent $\Rightarrow$ Admissible

(a) Given that a heuristic h is such that $h(G) = 0$, where G is any goal state, prove that if h is consistent, then it must be admissible.



- Let $v_1 \to v_2 \to \cdots \to v_n$ be the optimal path from v_1 to G (with minimum cost).
- We have $h(v_{i-1}) - h(v_i) \leq c(v_{i-1}, a_{i-1}, v_i)$ by consistency of h.
- Summing up over all i, we have

$$(h(v_1) \overbrace{- h(v_2)) + (h(v_2) - h(v_3)) + \cdots + (h(v_{n-1})}^{\text{cancels out}} - h(v_n)) \leq \overbrace{\sum c(v_{i-1}, a_{i-1}, v_i)}^{\text{sum of all costs}}$$

$$h(v_1) - h(v_n) \leq h^*(v_1)$$

- Since $h(G) = 0$, we have $h(v_1) \leq h^*(v_1) \Rightarrow h$ is admissible.

Q3. Consistent $\Rightarrow$ Admissible

(b) Give an example of an admissible heurstic function that is not consistent.

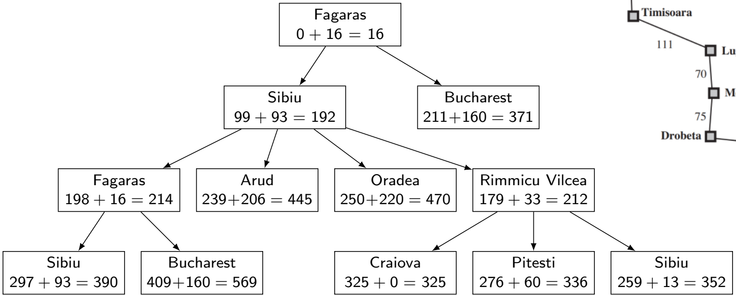
$$v \xrightarrow{1} u \xrightarrow{2} G$$

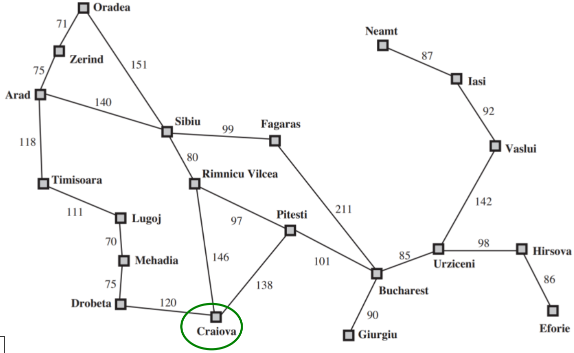
$$\begin{array}{cccc} h^*(n) & 3 & 2 & 0 \\ h(n) & 3 \xrightarrow{2} & 1 & 0 \end{array}$$

overestimate!

Q4. Fagaras to Craiova

(a) Trace A* search with SEARCH using the heuristic $h(n) = |h_{SLD}(\text{Craiov}) - h_{SLD}(n)|$.



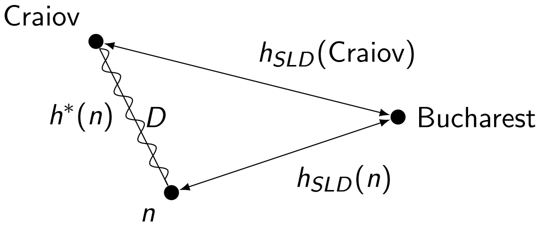


Values of h_{SLD} - straight-line distances to Bucharest

Arad	366	Mehadia	241
Bucharest	0	Neamt	234
Craiova	160	Oradea	380
Drobeta	242	Pitesti	100
Eforie	161	Rimnicu Vilcea	193
Fagaras	176	Sibiu	253
Giurgiu	77	Timisoara	329
Hirsova	151	Urziceni	80
Iasi	226	Vaslui	199
Lugoj	244	Zerind	374

Q4. Fagaras to Craiova

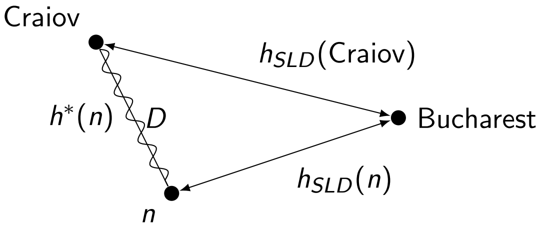
(b) Prove that $h(n) = |h_{SLD}(\text{Craiov}) - h_{SLD}(n)|$ is an admissible heuristic.



- Intuition: $h(n) \leq D \leq h^*(n)$.
- We want to show $D \geq h_{SLD}(\text{Craiov}) - h_{SLD}(n)$ and $D \geq h_{SLD}(n) - h_{SLD}(\text{Craiov})$.
- Changing terms: $D + h_{SLD}(n) \geq h_{SLD}(\text{Craiov})$ and $D + h_{SLD}(\text{Craiov}) \geq h_{SLD}(n)$.
  Triangle inequality!

Q4. Fagaras to Craiova

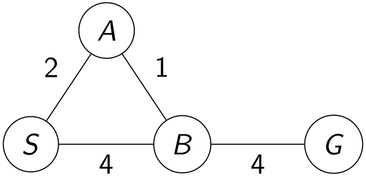
Proof:



- By triangle inequality, $D + h_{SLD}(n) \geq h_{SLD}(\text{Craiov})$ and $D + h_{SLD}(\text{Craiov}) \geq h_{SLD}(n)$.
- Therefore, $D \geq h_{SLD}(\text{Craiov}) - h_{SLD}(n)$ and $D \geq h_{SLD}(n) - h_{SLD}(\text{Craiov})$, i.e. $D \geq |h_{SLD}(\text{Craiov}) - h_{SLD}(n)|$.
- Since $D \leq h^*(n)$, we have $h(n) \leq D \leq h^*(n)$.

Q5. Inconsistent Heuristic

Show that A* using `SEARCH WITH VISITED MEMORY` returns a nonoptimal solution path when using an admissible but inconsistent $h(n)$. Then, show that `SEARCH` will return the optimal solution with the same heuristic.



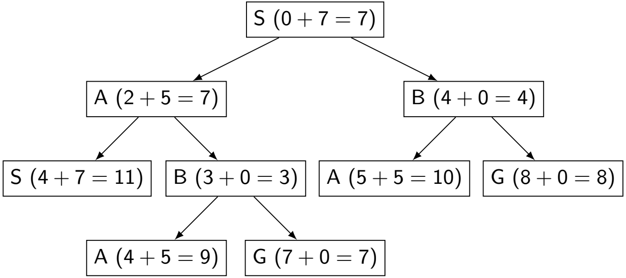
- The optimal solution path is $S \to A \to B \to G$.
- We want to trick A* to pick $S \to B \to G$ instead (i.e. B expanded before A).
  $\Rightarrow$ Make $h(A)$ large but $h(B)$ small.

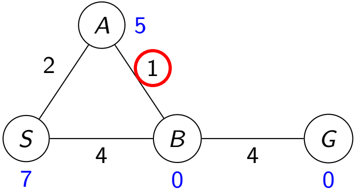
Q5. Inconsistent Heuristic

SEARCH WITH VISITED MEMORY:

- S expanded, $(A, 2+5=7)$ and $(B, 4+0=4)$ pushed into frontier.
- B expanded, $(A, 5+5=10)$ and $(G, 8+0=8)$ pushed.
- A expanded, $(S, 4+7=9)$ and $(B, 3+0=3)$ pushed.
- We do not revisit B (even though the cost is just 3). We reached G with cost 8. **Boom!**

SEARCH:





Bonus. A* Search vs Dijkstra

Extra Slide

Give a test case where Dijkstra algorithm fails to find the optimal path while A* search always finds the optimal path. Explain why that happens.

Bonus. A* Search vs Dijkstra

Extra Slide

- Dijkstra algorithm prioritizes the nodes by distance from source. Therefore, it visits G (distance 1) before A (distance 2).
- For A* search, any admissible heuristic must satisfy $h(A) \leq -2$. (This is "additional information" that A might lead a shorter path to G.) We always explore A first since $f(A) \leq 0 < 1 = f(G)$.

This is why A* search works even when there are negative edge weights.