Heuristic Search

Wenhu Chen

Lecture 18

Outline

Learning Goals

Why Heuristic Search

LCFS, GBFS, and A*

- Lowest-Cost-First Search
- Greedy Best-First Search
- A* Search

Designing an Admissible Heuristic

Pruning the Search Space

Learning goals

- Describe motivations for applying heuristic search algorithms.
- Trace the execution of and implement the Lowest-cost-first search, Greedy best-first search and A* search algorithm.
- Describe properties of the Lowest-cost-first, Greedy best-first and A* search algorithms.
- Design an admissible heuristic function for a search problem. Describe strategies for choosing among multiple heuristic functions.
- Describes strategies for pruning a search space.

Learning Goals

Why Heuristic Search

LCFS, GBFS, and A*

Designing an Admissible Heuristic

Pruning the Search Space

Why Heuristic Search?

How would ___ choose which one of the two states to expand?

- an uninformed search algorithm
- humans

5	3	
8	7	6
2	4	1

1	2	3
4	5	
7	8	6

Why Heuristic Search

An uninformed search algorithm

- considers every state to be the same.
- does not know which state is closer to the goal.
- may not find the optimal solution.

An heuristic search algorithm

- uses heuristics to estimate how close the state is to a goal.
- try to find the optimal solution.

Learning Goals

Why Heuristic Search

LCFS, GBFS, and A*

Lowest-Cost-First Search

Greedy Best-First Search

A* Search

Designing an Admissible Heuristic

Pruning the Search Space

The Cost Function

Suppose that we are executing a search algorithm
and we have added a path ending at n to the frontier.

$cost(n)$ is the actual cost of the path ending at n.

The Heuristic Function

Definition (search heuristic)

A **search heuristic** $h(n)$ is an estimate of the cost of the cheapest path from node n to a goal node.

In general, $h(n)$ can be arbitrary.

However, a good heuristic function has the following properties.

- problem-specific.
- non-negative.
- $h(n) = 0$ if n is a goal node.
- $h(n)$ must be easy to compute (without search).

LCFS, GBFS, and A*

- LCFS: remove the path with the lowest cost $cost(n)$.
- GBFS: remove the path with the lowest heuristic value $h(n)$.
- A*: remove the path with the lowest cost + heuristic value $cost(n) + h(n)$.

Learning Goals

Why Heuristic Search

LCFS, GBFS, and A*

- Lowest-Cost-First Search
- Greedy Best-First Search
- A* Search

Designing an Admissible Heuristic

Pruning the Search Space

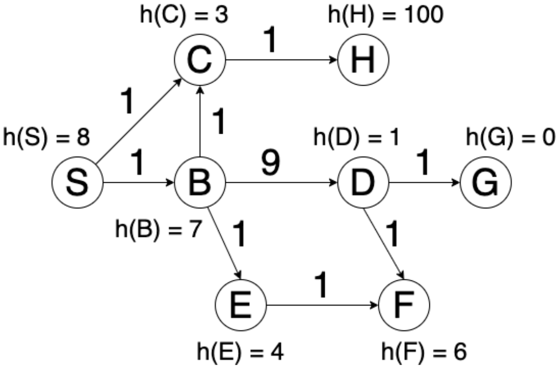
Lowest-cost-first search

- Frontier is a priority queue ordered by $cost(n)$.
- Expand the path with the lowest $cost(n)$.

$\rightarrow$ a.k.a. Dijkstra's shortest path algorithm.

Trace LCFS on a search graph

If there is a tie, remove nodes from the frontier in alphabetical order.

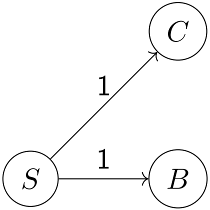


Trace LCFS on a search graph

Frontier: (S)

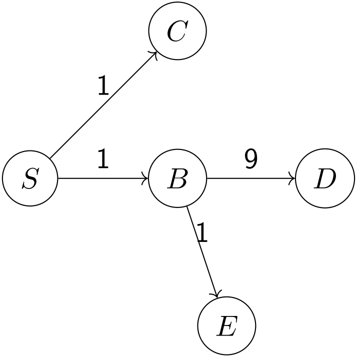
Trace LCFS on a search graph

Frontier: (S: 0) $\rightarrow$ (SB: 1, SC: 1)



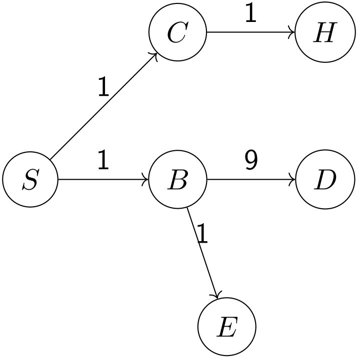
Trace LCFS on a search graph

Frontier: (SB: 1, SC: 1) $\rightarrow$ (SC: 1, SBE: 2, SBD: 10)



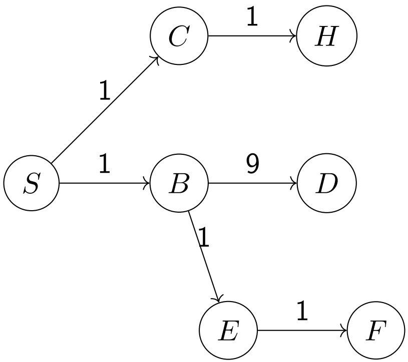
Trace LCFS on a search graph

Frontier: (SC: 1, SBE: 2, SBD: 10) $\rightarrow$ (SBE: 2, SCH: 2, SBD: 10)



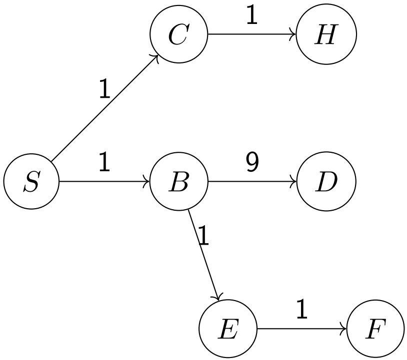
Trace LCFS on a search graph

Frontier: (SBE: 2, SCH: 2, SBD: 10) → (SCH: 2, SBEF: 3, SBD: 10)



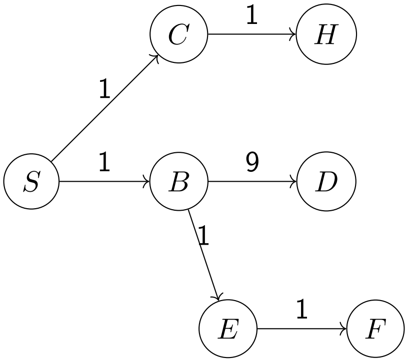
Trace LCFS on a search graph

Frontier: (SCH: 2, SBEF: 3, SBD: 10) $\rightarrow$ (SBEF: 3, SBD: 10)



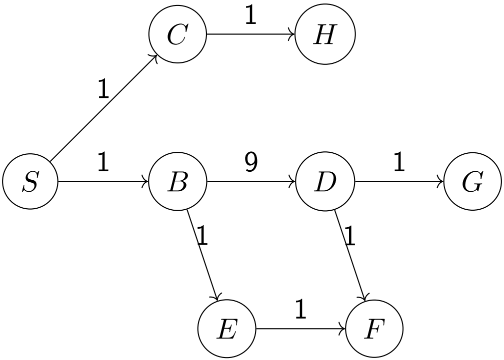
Trace LCFS on a search graph

Frontier: (SBEF: 3, SBD: 10) $\rightarrow$ (SBD: 10)



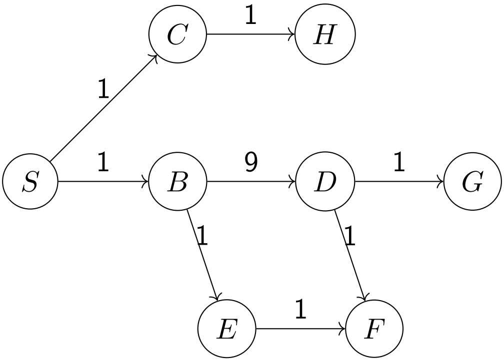
Trace LCFS on a search graph

Frontier: (SBD: 10) $\rightarrow$ (SBDF: 11, SBDG: 11)



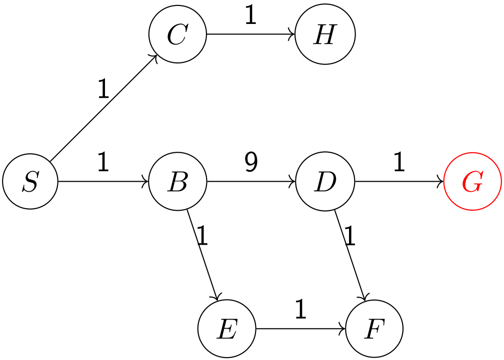
Trace LCFS on a search graph

Frontier: (SBDF: 11, SBDG: 11) $\rightarrow$ (SBDG: 11)



Trace LCFS on a search graph

Frontier: (SBDG: 11) $\rightarrow$ ()



Properties of LCFS

- Space and Time Complexities

Properties of LCFS

- Space and Time Complexities

  Both complexities are exponential.
  LCFS examines a lot of paths to ensure that
  it returns the optimal solution first.

Properties of LCFS

- Space and Time Complexities

  Both complexities are exponential.
  LCFS examines a lot of paths to ensure that
  it returns the optimal solution first.

- Completeness and Optimality

Properties of LCFS

- Space and Time Complexities

  Both complexities are exponential.
  LCFS examines a lot of paths to ensure that
  it returns the optimal solution first.

- Completeness and Optimality

  Yes and yes under mild conditions:
  (1) The branching factor is finite.
  (2) The cost of every edge is bounded below by a positive constant.

Learning Goals

Why Heuristic Search

LCFS, GBFS, and A*

Lowest-Cost-First Search

Greedy Best-First Search

A* Search

Designing an Admissible Heuristic

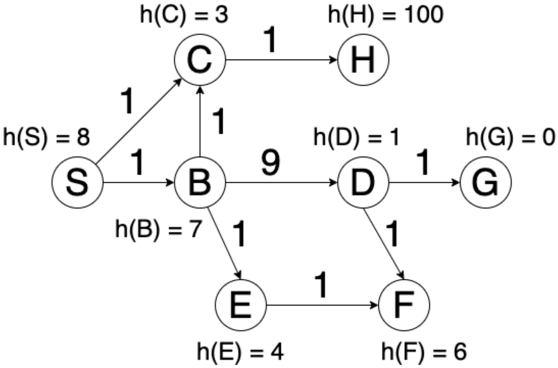
Pruning the Search Space

Greedy Best-First Search

- Frontier is a priority queue ordered by $h(n)$.
- Expand the node with the lowest $h(n)$.

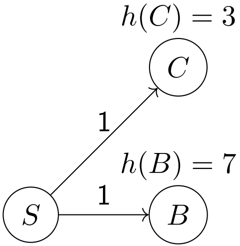
Trace GBFS on a search graph

If there is a tie, remove nodes from the frontier in alphabetical order.



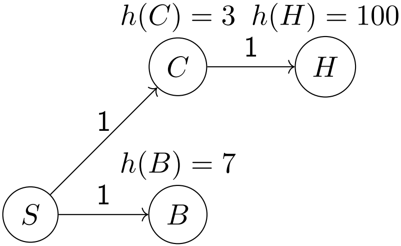
Trace GBFS on a search graph

Frontier: (S) $\rightarrow$ (SC: 3, SB: 7)



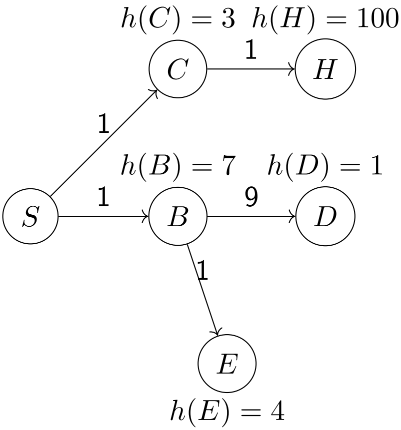
Trace GBFS on a search graph

Frontier: (SC: 3, SB: 7) $\rightarrow$ (SB: 7, SCH: 100)



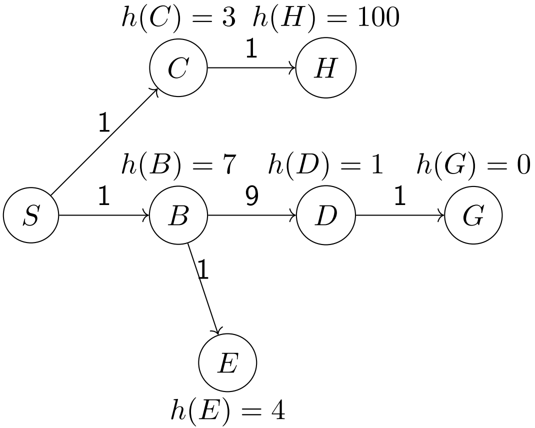
Trace GBFS on a search graph

Frontier: (SB: 7, SCH: 100) $\rightarrow$ (SBD: 1, SBE: 4, SCH: 100)



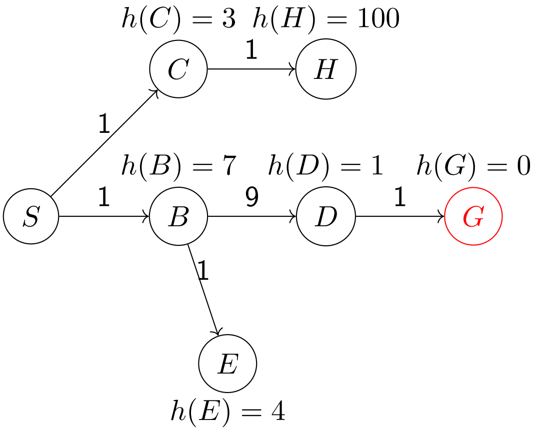
Trace GBFS on a search graph

Frontier: (SBD: 1, SBE: 4, SCH: 100) $\rightarrow$ (SBDG: 0, SBE: 4, SCH: 100)



Trace GBFS on a search graph

Frontier: (SBDG: 0, SBE: 4, SCH: 100) $\rightarrow$ (SBE: 4, SCH: 100)



Properties of GBFS

- Space and Time Complexities

Properties of GBFS

- Space and Time Complexities

  Both complexities are exponential.

Properties of GBFS

- Space and Time Complexities

  Both complexities are exponential.

- Completeness and Optimality

Properties of GBFS

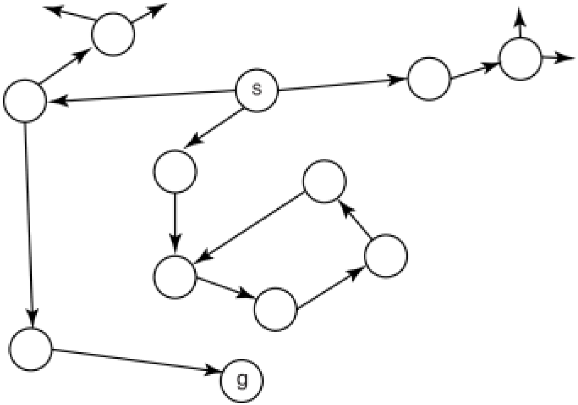
- Space and Time Complexities

  Both complexities are exponential.

- Completeness and Optimality

  No, GBFS is not complete. It could be stuck in a cycle.
  No, GBFS is not optimal. GBFS may return a sub-optimal path first.

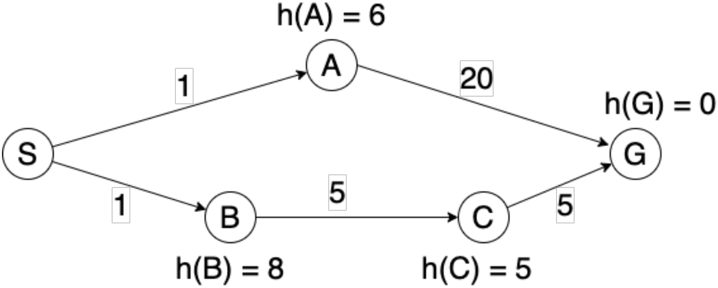
Greedy BFS: will it find a solution/terminate?



$\rightarrow$ The cost of an arc is its length.

The heuristic function is the Euclidean straight line distance.

Greedy BFS: will it find the optimal solution?



$\rightarrow$ Path found by Greedy BFS: $S \rightarrow A \rightarrow G$, cost $= 21$.

The optimal solution: $S \rightarrow B \rightarrow C \rightarrow G$, cost $= 11$.

Learning Goals

Why Heuristic Search

LCFS, GBFS, and A*

Lowest-Cost-First Search

Greedy Best-First Search

A* Search

Designing an Admissible Heuristic

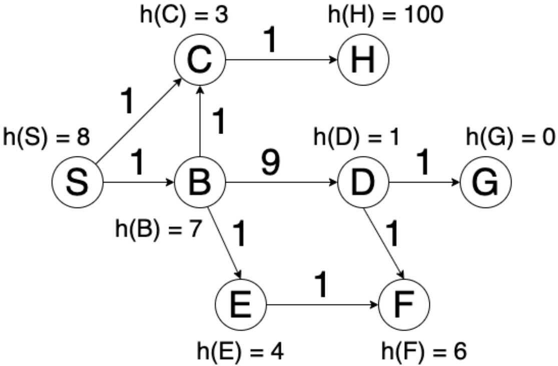
Pruning the Search Space

A* Search

- Frontier is a priority queue ordered by $f(n) = cost(n) + h(n)$.
- Expand the node with the lowest $f(n)$.

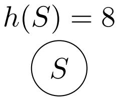
Trace A* search on a search graph

If there is a tie, remove nodes from the frontier in alphabetical order.



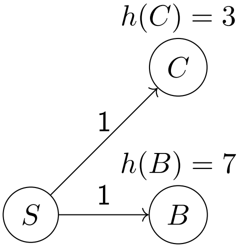
Trace A* on a search graph

Frontier: (S: 8)



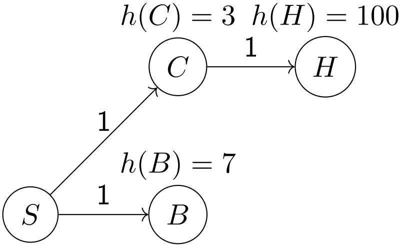
Trace A* on a search graph

Frontier: (S: 8) $\rightarrow$ (SC: 4, SB: 8)



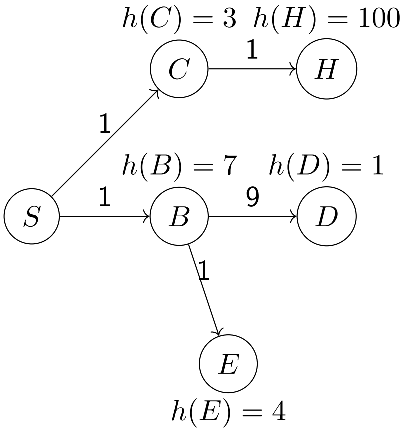
Trace A* on a search graph

Frontier: (SC: 4, SB: 8) $\rightarrow$ (SB: 8, SCH: 102)



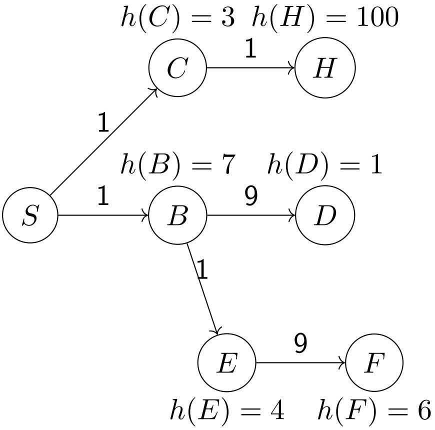
Trace A* on a search graph

Frontier: (SB: 8, SCH: 102) $\rightarrow$ (SBE: 6, SBD:11, SCH: 102)



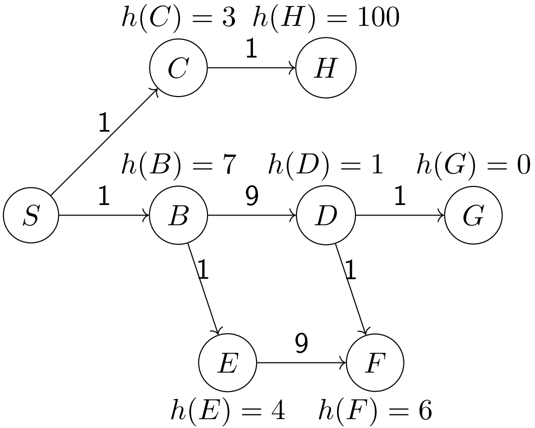
Trace A* on a search graph

Frontier: (SBE: 6, SBD:11, SCH: 102) $\rightarrow$ (SBD:11, SBEF: 17, SCH: 102)



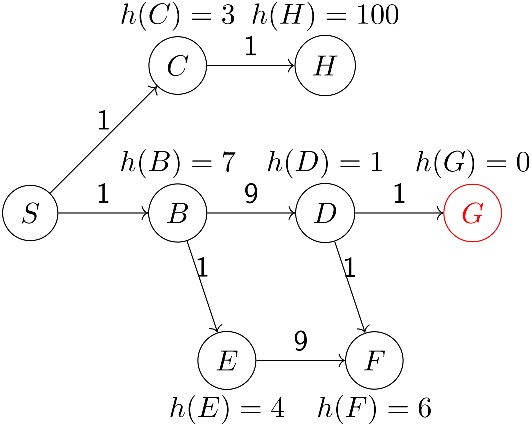
Trace A* on a search graph

Frontier: (SBD:11, SBEF: 17, SCH: 102) → (SBDG: 11, SBDF: 17, SBEF: 17, SCH: 102)



Trace A* on a search graph

Frontier: (SBDG: 11, SBDF: 17, SBEF: 17, SCH: 102) → (SBDF: 17, SBEF: 17, SCH: 102)



Properties of A* Search

- Space and Time Complexities

Properties of A* Search

- Space and Time Complexities

  Both complexities are exponential.

Properties of A* Search

- Space and Time Complexities

  Both complexities are exponential.

- Completeness and Optimality

Properties of A* Search

- Space and Time Complexities

  Both complexities are exponential.

- Completeness and Optimality

  Yes and Yes, given mild conditions on the heuristic function.

A* is Optimal

Definition (admissible heuristic)

A heuristic $h(n)$ is admissible if it never over-estimates the cost of the cheapest path from node n to a goal node.

Theorem (Optimality of A*)

If the heuristic $h(n)$ is admissible, the solution found by A is optimal.*

A* is Optimal

- Assuming you have many paths in the frontier: $(S \to G : C^*, \cdots, S \to N : C^n)$, and $C^* \leq C^n$.
- If there a path through N to G has a lower cost of $C' < C^*$.
- According to admissibility, $C^n \leq C' < C^*$.
- It's contradictory to our assumption.

A* is Optimally Efficient

Among all optimal algorithms that start from the same start node and use the same heuristic, A* expands the fewest nodes.

A* is Optimally Efficient

Among all optimal algorithms that start from the same start node and use the same heuristic, A* expands the fewest nodes.

$\rightarrow$ No algorithm with the same information can do better.

A* expands the minimum number of nodes to find the optimal solution.

Intuition for a proof: any algorithm that does not expand all nodes with $f(n) < C^*$ run the risk of missing the optimal solution.

Learning Goals

Why Heuristic Search

LCFS, GBFS, and A*

Designing an Admissible Heuristic

Pruning the Search Space

Some Heuristic Functions for 8-Puzzle

- *Manhattan Distance Heuristic:*

  The sum of the Manhattan distances of the tiles from their goal positions

- *Misplaced Tile Heuristic:*

  The number of tiles that are NOT in their goal positions

Both heuristic functions are admissible.

Initial State

5	3	
8	7	6
2	4	1

Goal State

1	2	3
4	5	6
7	8	

Constructing an Admissible Heuristic

1. Define a relaxed problem by simplifying or removing constraints on the original problem.
2. Solve the relaxed problem without search.
3. The cost of the optimal solution to the relaxed problem is an admissible heuristic for the original problem.

$\rightarrow$ Simplifying or removing constraints — making the problem easier.

For an easier problem, the cost of the optimal solution should be smaller than that of the original problem.

Constructing an Admissible Heuristic for 8-Puzzle

8-puzzle: A tile can move from square A to square B

- if A and B are adjacent, and
- B is empty.

Which heuristics can we derive from relaxed versions of this problem?

Q: Constructing an Admissible Heuristic

Q #1: Which heuristics can we derive from the following relaxed 8-puzzle problem?

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

(A) The Manhattan distance heuristic

(B) The Misplaced tile heuristic

(C) Another heuristic not described above

Q: Constructing an Admissible Heuristic

Q #1: Which heuristics can we derive from the following relaxed 8-puzzle problem?

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

(A) **The Manhattan distance heuristic**

(B) The Misplaced tile heuristic

(C) Another heuristic not described above

$\rightarrow$ (A) is correct

Desirable Heuristic Properties

- We want a heuristic to be admissible.

  $\rightarrow$ A* is optimal.

- Want a heuristic to have higher values (close to h^*).

  $\rightarrow$ The closer h is to h^*, the most accurate h is.

- Prefer a heuristic that is very different for different states.

  $\rightarrow$ h should help us choose among different paths. If h is close to constant, not useful.

Dominating Heuristic

Definition (dominating heuristic)

Given heuristics $h_1(n)$ and $h_2(n)$. $h_2(n)$ dominates $h_1(n)$ if

- $(\forall n \ (h_2(n) \geq h_1(n)))$.
- $(\exists n \ (h_2(n) > h_1(n)))$.

Theorem

If $h_2(n)$ dominates $h_1(n)$, A using h_2 will never expand more nodes than A* using h_1.*

Q: Which Heuristic of 8-puzzle is Better?

Q #2: Which of the two heuristics of the 8-puzzle is better?

(A) The Manhattan distance heuristic dominates the Misplaced tile heuristic.

(B) The Misplaced tile heuristic dominates the Manhattan distance heuristic.

(C) Neither dominates the other one.

Q: Which Heuristic of 8-puzzle is Better?

Q #2: Which of the two heuristics of the 8-puzzle is better?

(A) **The Manhattan distance heuristic dominates the Misplaced tile heuristic.**

(B) The Misplaced tile heuristic dominates the Manhattan distance heuristic.

(C) Neither dominates the other one.

$\rightarrow$ If a tile is out of place, Misplaced tile will $+1$. Manhattan distance will add at least 1 and maybe more. So Manhattan distance heuristic always gives a larger value.

Learning Goals

Why Heuristic Search

LCFS, GBFS, and A*

Designing an Admissible Heuristic

Pruning the Search Space

Cycle Pruning

- What is cycle pruning?

Cycle Pruning

- What is cycle pruning?

  Whenever we realize that we are following a cycle, we should stop expanding the path.

Cycle Pruning

- What is cycle pruning?

  Whenever we realize that we are following a cycle, we should stop expanding the path.

- Why do we want to perform cycle pruning?

Cycle Pruning

- What is cycle pruning?

  Whenever we realize that we are following a cycle, we should stop expanding the path.

- Why do we want to perform cycle pruning?

  Cycles may cause an algorithm to not terminate, e.g. DFS. Exploring a cycle is a waste of time since it cannot be part of a solution.

Search w/ Cycle Pruning

- How do we perform cycle pruning?

Search w/ Cycle Pruning

- How do we perform cycle pruning?

Algorithm 2 Search w/ Cycle Pruning

...
for every neighbour n of n_k **do**
  if $n \notin \langle n_0, \ldots, n_k \rangle$ **then**
    add $\langle n_0, \ldots, n_k, n \rangle$ to frontier;
...

Search w/ Cycle Pruning

- How do we perform cycle pruning?

Algorithm 3 Search w/ Cycle Pruning

```
...
for every neighbour n of n_k do
    if n ∉ ⟨n_0, ..., n_k⟩ then
        add ⟨n_0, ..., n_k, n⟩ to frontier;
...
```

- What is the complexity of cycle pruning for DFS and BFS?

Search w/ Cycle Pruning

- How do we perform cycle pruning?

Algorithm 4 Search w/ Cycle Pruning

...
for every neighbour n of n_k **do**
    if $n \notin \langle n_0, \ldots, n_k \rangle$ **then**
        add $\langle n_0, \ldots, n_k, n \rangle$ to frontier;
...

- What is the complexity of cycle pruning for DFS and BFS?

Time complexity: linear to the path length.

Multiple-Path Pruning

- Why do we want to perform multi-path pruning?

Multiple-Path Pruning

- Why do we want to perform multi-path pruning?

  If we have already found a path to a node, we can discard other paths to the same node.

Multiple-Path Pruning

- Why do we want to perform multi-path pruning?

  If we have already found a path to a node, we can discard other paths to the same node.

- What is the relationship between cycle pruning and multi-path pruning?

Multiple-Path Pruning

- Why do we want to perform multi-path pruning?

  If we have already found a path to a node, we can discard other paths to the same node.

- What is the relationship between cycle pruning and multi-path pruning?

  Cycle pruning is a special case of multi-path pruning.

Search w/ Multi-Path Pruning

How do we perform multi-path pruning?

Search w/ Multi-Path Pruning

How do we perform multi-path pruning?

Algorithm 6 Search w/ Multi-Path Pruning

procedure $\textsc{Search}$(Graph, Start node s, Goal test $goal(n)$)
  frontier := $\{\langle s\rangle\}$;
  explored := $\{\}$;
  while frontier is not empty **do**
    select and **remove** path $\langle n_0, \ldots, n_k\rangle$ from frontier;
    if $n_k \notin$ **explored then**
      add n_k **to explored**
      if goal(n_k) **then**
        return $\langle n_0, \ldots, n_k\rangle$;
      for every neighbour n of n_k **do**
        add $\langle n_0, \ldots, n_k, n\rangle$ to frontier;
  return no solution

Search w/ Multi-Path Pruning

There are some caveats:

- Node will be added to the 'explored' set once it's **explored**
- The longer paths leading to 'explored' set will **still be added to frontier**, they are just not explored.
- It saves computation but increases space consumption.

A problem with multi-path pruning

- Multi-path pruning says that we keep the first path to a node and discard the rest.
- What if the first path to a node is not the least-cost path?
- Can multi-path pruning cause a search algorithm to fail to find the optimal solution?

Lowest-cost-first search w/ multi-path pruning

Can Lowest-Cost-First Search with multi-path pruning discard the optimal solution?

(A) Yes, it is possible.

(B) No, it is not possible.

Lowest-cost-first search w/ multi-path pruning

Can Lowest-Cost-First Search with multi-path pruning discard the optimal solution?

(A) Yes, it is possible.

(B) **No, it is not possible.**

→ (B) No, it is not possible.
LCFS always finds the least-cost path first.

A* search w/ multi-path pruning

Can A* search with an admissible heuristic and multi-path pruning discard the optimal solution?

(A) Yes, it is possible.

(B) No, it is not possible.

A* search w/ multi-path pruning

Can A* search with an admissible heuristic and multi-path pruning discard the optimal solution?

(A) **Yes, it is possible.**

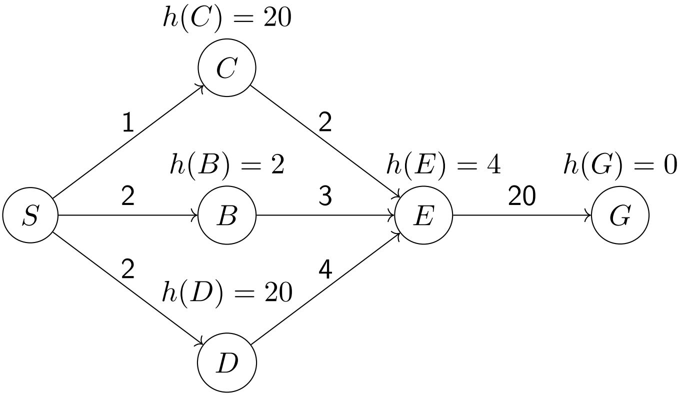
(B) No, it is not possible.

$\rightarrow$ (A) Yes, it is possible.

When we select a path to a node for the first time,
this path may not be the least-cost path to the node.

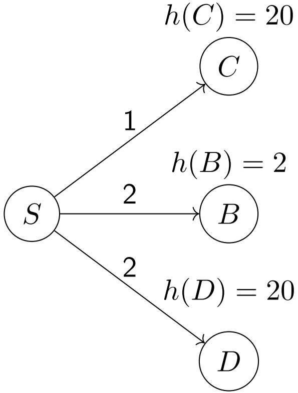
A* with multi-path pruning is not optimal.

A* search w/ multi-path pruning



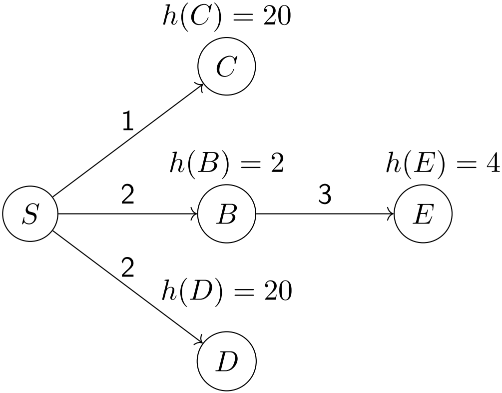
A* search w/ multi-path pruning

Frontier: (S) $\rightarrow$ (SB: 4, SC: 21, SD: 22)
Explored: (S)



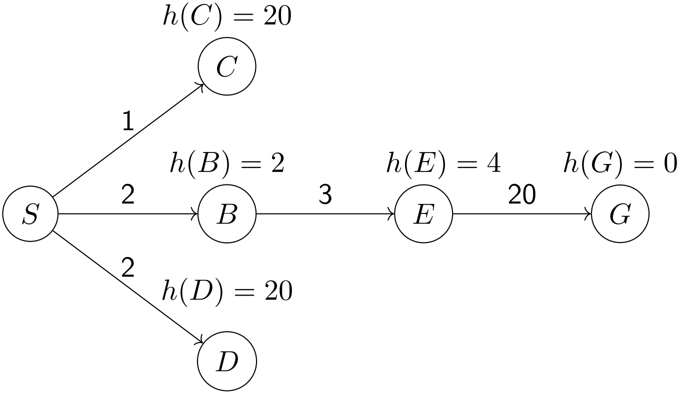
A* search w/ multi-path pruning

Frontier: (SB: 4, SC: 21, SD: 22) → (SBE: 9, SC: 21, SD: 22)
Explored: (S, B)



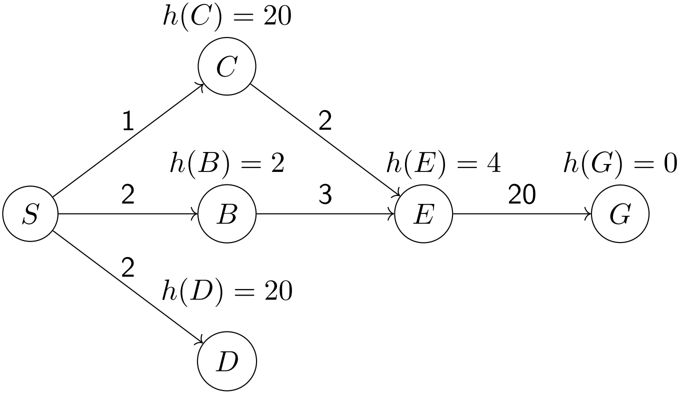
A* search w/ multi-path pruning

Frontier: (SBE: 9, SC: 21, SD: 22) → (SC: 21, SD: 22, SBEG: 25)
Explored: (S, B, E)



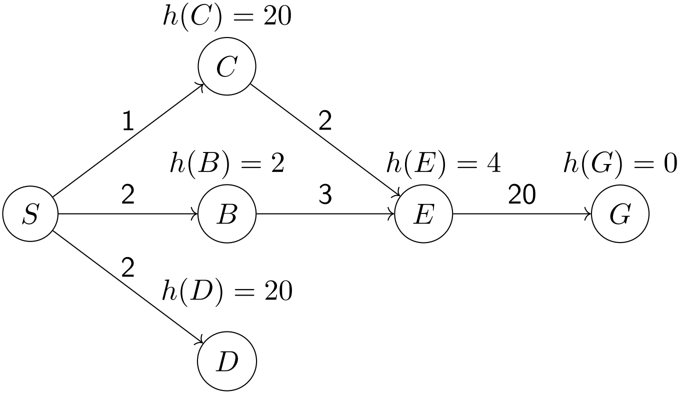
A* search w/ multi-path pruning

Frontier: (SBE: 9, SC: 21, SD: 22) → (SCE: 7, SD: 22, SBEG: 25)
Explored: (S, B, E)



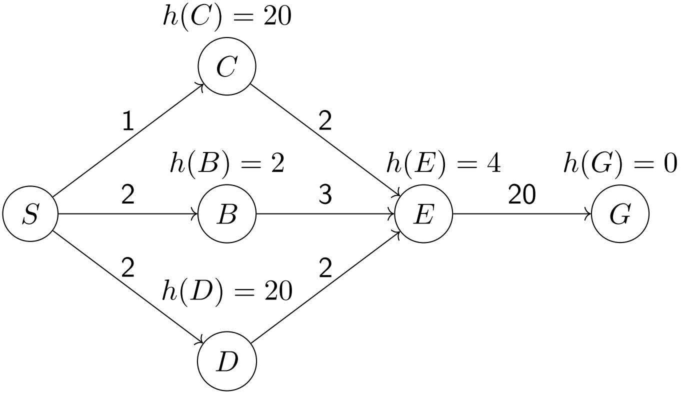
A* search w/ multi-path pruning

Frontier: (SCE: 7, SD: 22, SBEG: 25) $\rightarrow$ (SD: 22, SBEG: 25)
Explored: (S, B, E)



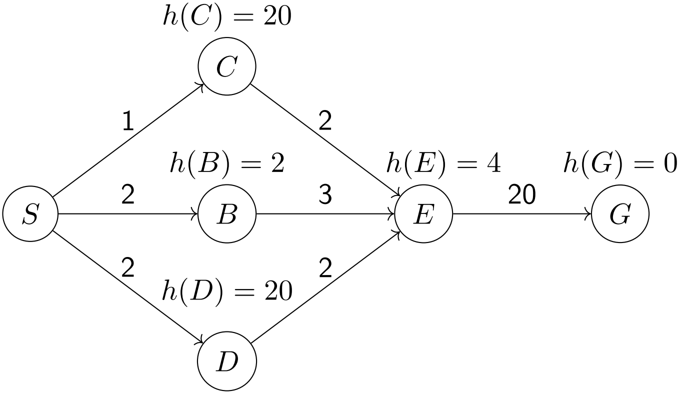
A* search w/ multi-path pruning

Frontier: (SD: 22, SBEG: 25) $\rightarrow$ (SDE: 8, SBEG: 25)
Explored: (S, B, E)



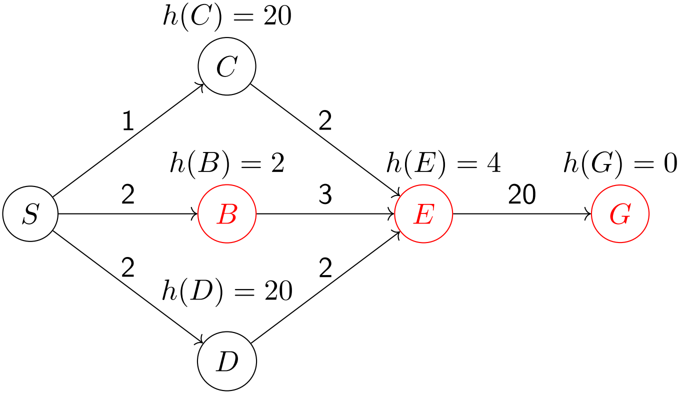
A* search w/ multi-path pruning

Frontier: (SDE: 8, SBEG: 25) → (SBEG: 25)
Explored: (S, B, E)



A* search w/ multi-path pruning

Frontier: (SBEG: 25) $\rightarrow$ ()
Explored: (S, B, E, G)



Finding optimal solution w/ multi-path pruning

What if a subsequent path to n is shorter than the first path found?

- Remove all paths from the frontier that use the longer path.
- Change the initial segment of the paths on the frontier to use the shorter path.
- Make sure that we find the least-cost path to a node first.

When the does multi-path pruning not work?

Assuming we have a frontier $(s \to n, \cdots, s \to n')$, and we are exploring node n.

- If there exists another path through n' to n with lower f-value.

When the does multi-path pruning not work?

Assuming we have a frontier $(s \to n, \cdots, s \to n')$, and we are exploring node n.

- If there exists another path through n' to n with lower f-value.
- 1) we have $h(n) + cost(n) > h(n) + cost(n') + cost(n', n)$, e.g. $cost(n) - cost(n') > cost(n', n)$

When the does multi-path pruning not work?

Assuming we have a frontier $(s \to n, \cdots, s \to n')$, and we are exploring node n.

- If there exists another path through n' to n with lower f-value.
- 1) we have $h(n) + cost(n) > h(n) + cost(n') + cost(n', n)$, e.g. $cost(n) - cost(n') > cost(n', n)$
- 2) node n is already explored, so $h(n) + cost(n) \leq h(n') + cost(n')$

When the does multi-path pruning not work?

Assuming we have a frontier $(s \rightarrow n, \cdots, s \rightarrow n')$, and we are exploring node n.

- If there exists another path through n' to n with lower f-value.
- 1) we have $h(n) + cost(n) > h(n) + cost(n') + cost(n', n)$, e.g. $cost(n) - cost(n') > cost(n', n)$
- 2) node n is already explored, so $h(n) + cost(n) \leq h(n') + cost(n')$
- Combine these two, we have $h(n') - h(n) \geq cost(n) - cost(n') > cost(n', n)$

When the does multi-path pruning not work?

Assuming we have a frontier $(s \to n, \cdots, s \to n')$, and we are exploring node n.

- If there exists another path through n' to n with lower f-value.
- 1) we have $h(n) + cost(n) > h(n) + cost(n') + cost(n', n)$, e.g. $cost(n) - cost(n') > cost(n', n)$
- 2) node n is already explored, so $h(n) + cost(n) \leq h(n') + cost(n')$
- Combine these two, we have $h(n') - h(n) \geq cost(n) - cost(n') > cost(n', n)$
- Such scenario only happens when there exists two nodes n and n' with $h(n') - h(n) > cost(n', n)$.

Consistent Heuristic

- An admissible heuristic requires that:
  For any node m and any goal node g,

$$h(m) - h(g) \leq cost(m, g).$$

Consistent Heuristic

- An admissible heuristic requires that:
  For any node m and any goal node g,

$$h(m) - h(g) \leq cost(m, g).$$

- To ensure that A* with multi-path pruning is optimal,
  we need a consistent heuristic function: For any two nodes m and n,

$$h(m) - h(n) \leq cost(m, n).$$

Consistent Heuristic

- An admissible heuristic requires that:
  For any node m and any goal node g,

  $$h(m) - h(g) \leq cost(m, g).$$

- To ensure that A* with multi-path pruning is optimal,
  we need a consistent heuristic function: For any two nodes m and n,

  $$h(m) - h(n) \leq cost(m, n).$$

- A consistent heuristic satisfies the monotone restriction:
  For any edge from m to n,

  $$h(m) - h(n) \leq cost(m, n).$$

Constructing Consistent Heuristics

- Most admissible heuristic functions are consistent.
- It's challenging to come up with a heuristic function that is admissible but not consistent.

Summary of Search Strategies

Strategy	Frontier Selection	Halts?	Space	Time
Depth-first	Last node added	No	Linear	Exp
Breadth-first	First node added	Yes	Exp	Exp
Lowest-cost-first	min $cost(n)$	Yes	Exp	Exp
Greedy Best-first	min $h(n)$	No	Exp	Exp
A*	min $cost(n) + h(n)$	Yes	Exp	Exp