CMSC 471
Artificial Intelligence

Search

KMA Solaiman – ksolaima@umbc.edu

Slide adoption courtesy: Tim Finin, Frank Ferraro (UMBC),
Charles R. Dyer (UW-Madison), Dan Klein and Pieter Abbeel (CS188, UC Berkeley)

A General Searching Algorithm

Core ideas:

1. Maintain a list of frontier (fringe) nodes
   1. Nodes coming *into* the frontier have been explored
   2. Nodes going out of the frontier have not been explored
2. Iteratively select nodes from the frontier and explore unexplored nodes from the frontier
3. Stop when you reach your **goal**

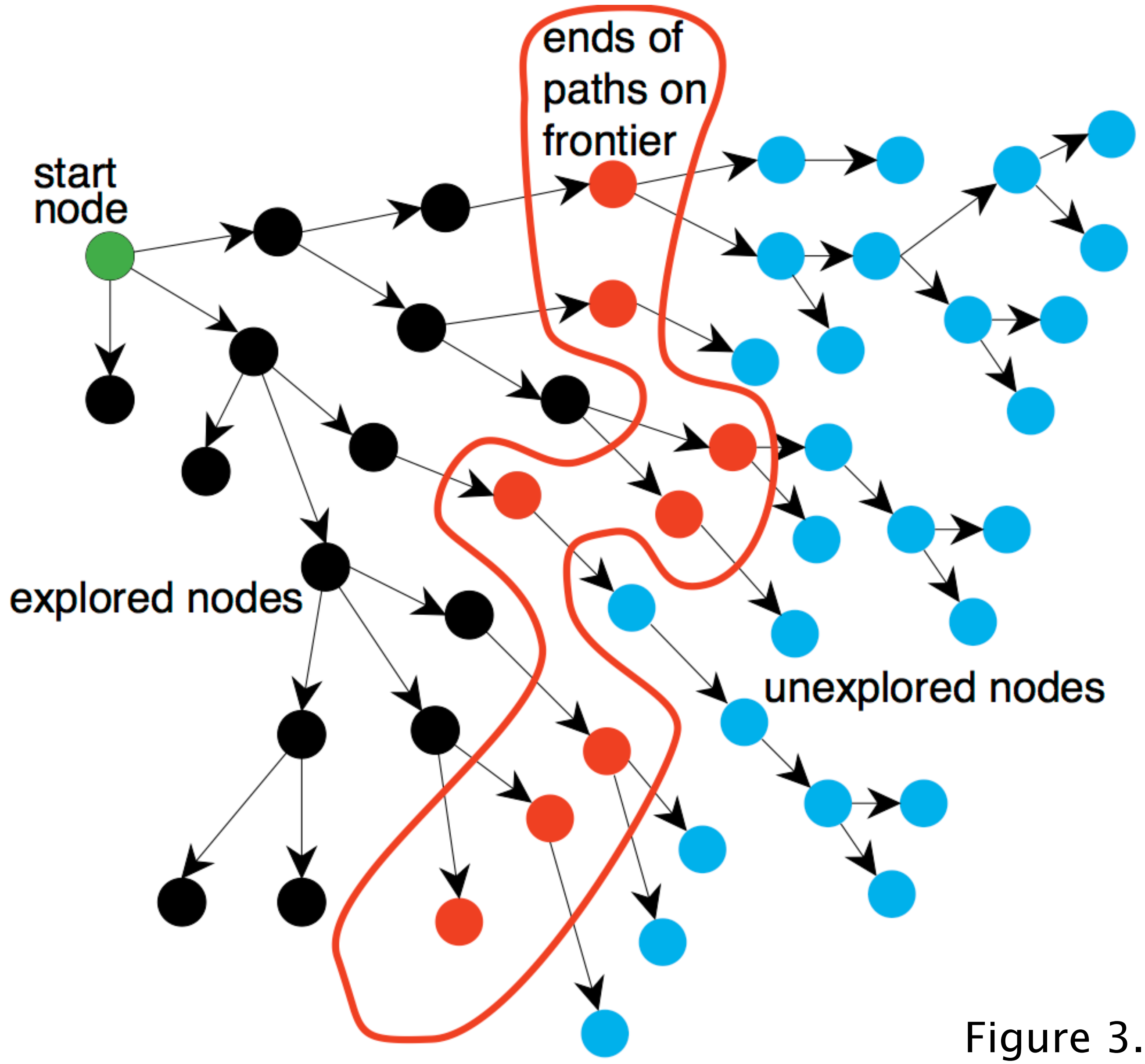


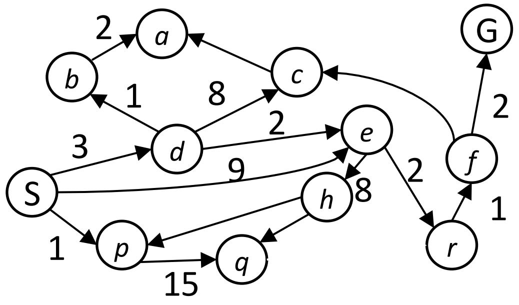
Figure 3.3

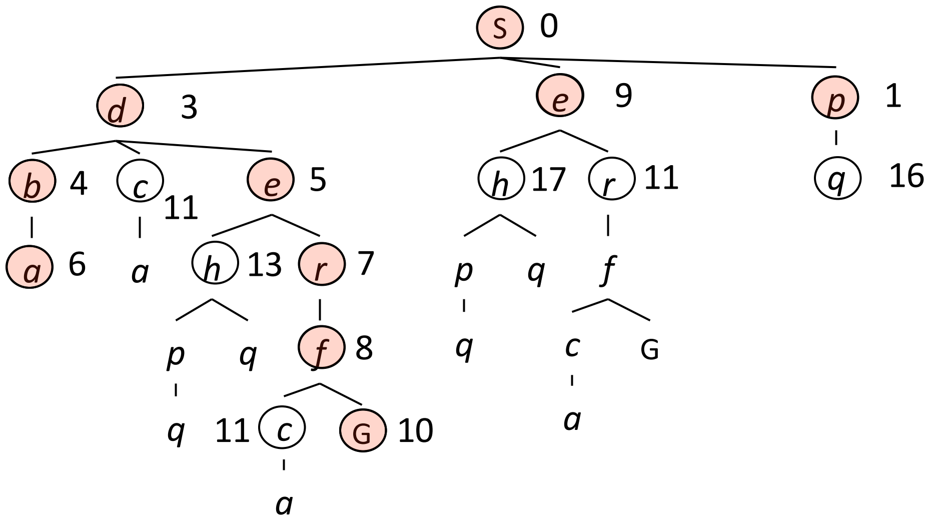
Uniform Cost Search

g(n) = cost from root to n

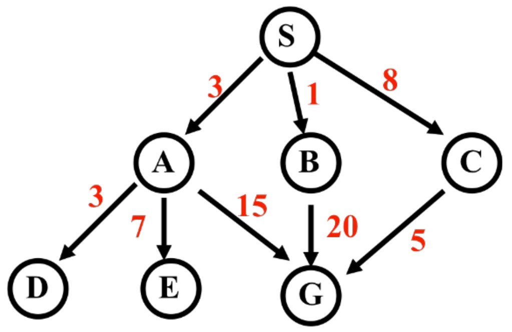
Strategy: expand lowest g(n)

Frontier is a priority queue sorted by g(n)





Uniform-Cost Search



Expanded node	Nodes list
	{ S^0 }
S^0	{ B^1 A^3 C^8 }
B^1	{ A^3 C^8 G^{21} }
A^3	{ D^6 C^8 E^{10} G^{18} G^{21} }
D^6	{ C^8 E^{10} G^{18} G^{21} }
C^8	{ E^{10} G^{13} G^{18} G^{21} }
E^{10}	{ G^{13} G^{18} G^{21} }
G^{13}	{ G^{18} G^{21} }

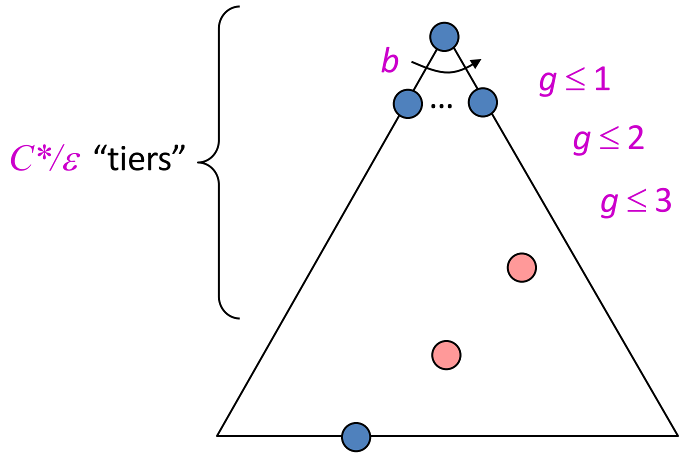
priority queue

Solution path found is S C G, cost 13

Number of nodes expanded (including goal node) = 7

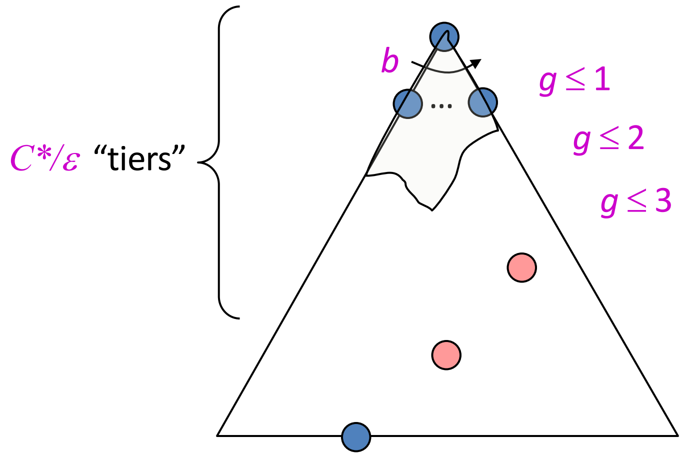
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?



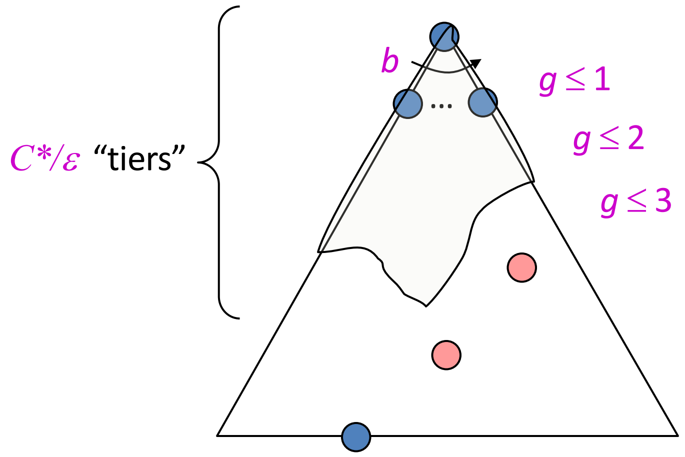
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?



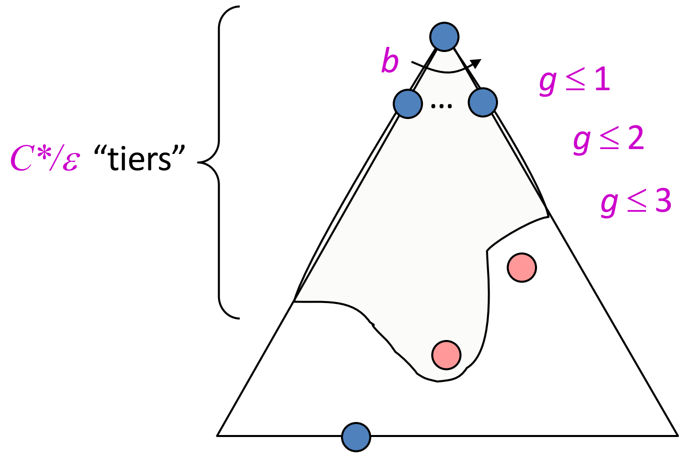
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?



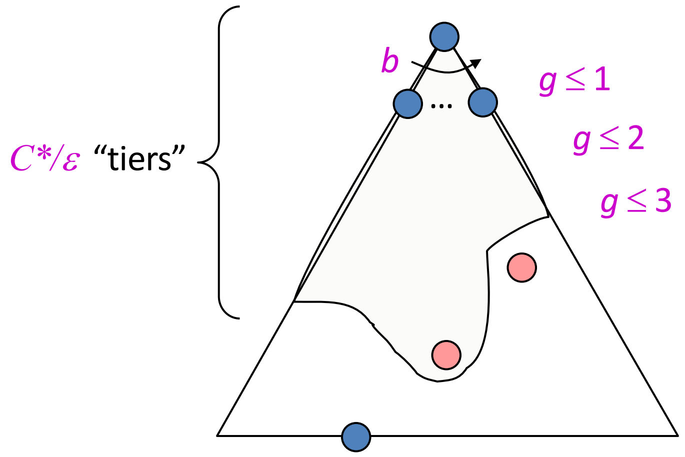
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?



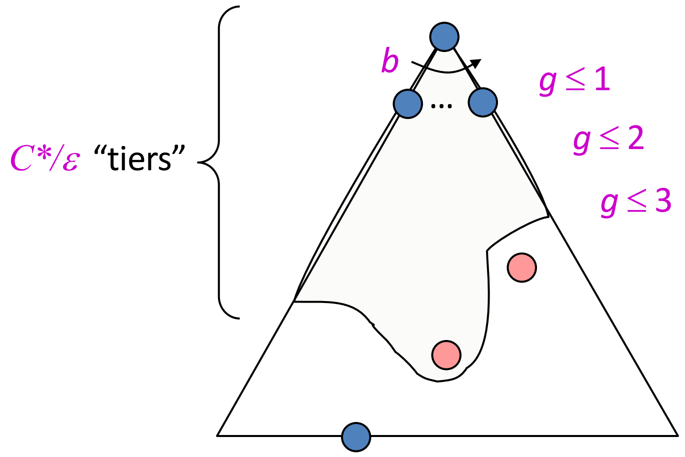
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?
  - Processes all nodes with cost less than cheapest solution!



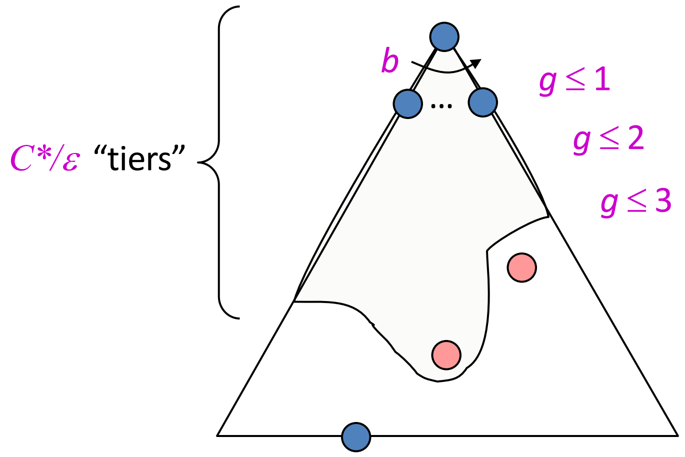
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?
  - Processes all nodes with cost less than cheapest solution!
  - If that solution costs C^* and arcs cost at least ε, then the "effective depth" is roughly C^*/ε



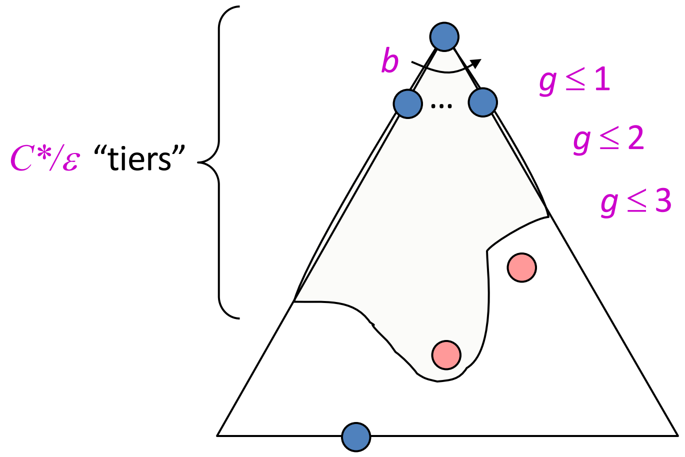
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?
  - Processes all nodes with cost less than cheapest solution!
  - If that solution costs C^* and arcs cost at least ε, then the "effective depth" is roughly C^*/ε
  - Takes time $O(b^{C^*/\varepsilon})$ (exponential in effective depth)



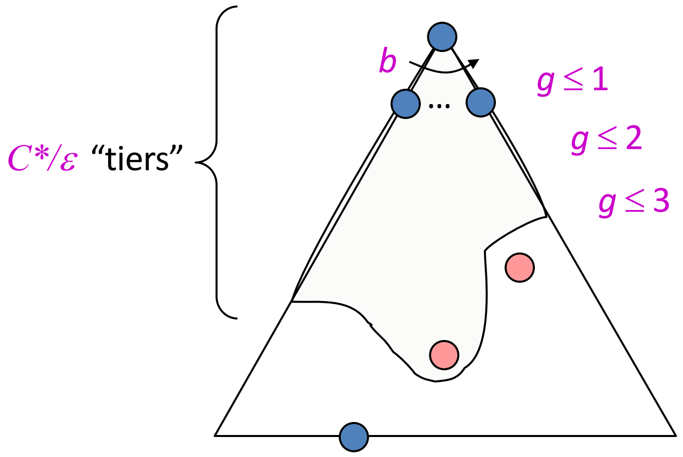
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?
  - Processes all nodes with cost less than cheapest solution!
  - If that solution costs C^* and arcs cost at least ε, then the "effective depth" is roughly C^*/ε
  - Takes time $O(b^{C^*/\varepsilon})$ (exponential in effective depth)
- How much space does the frontier take?



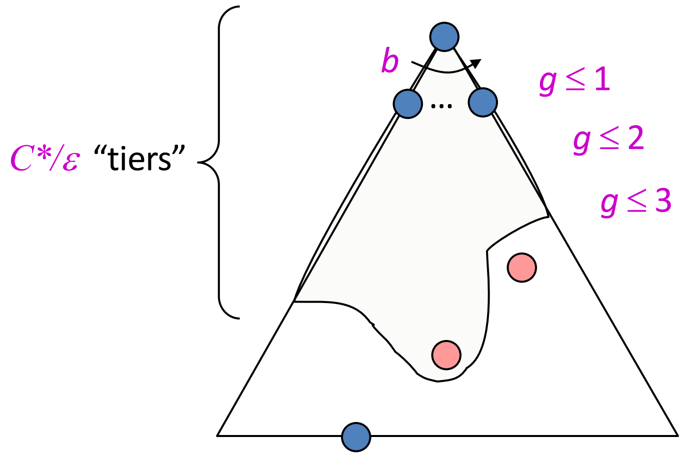
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?
  - Processes all nodes with cost less than cheapest solution!
  - If that solution costs C^* and arcs cost at least ε, then the "effective depth" is roughly C^*/ε
  - Takes time $O(b^{C^*/\varepsilon})$ (exponential in effective depth)
- How much space does the frontier take?
  - Has roughly the last tier, so $O(b^{C^*/\varepsilon})$



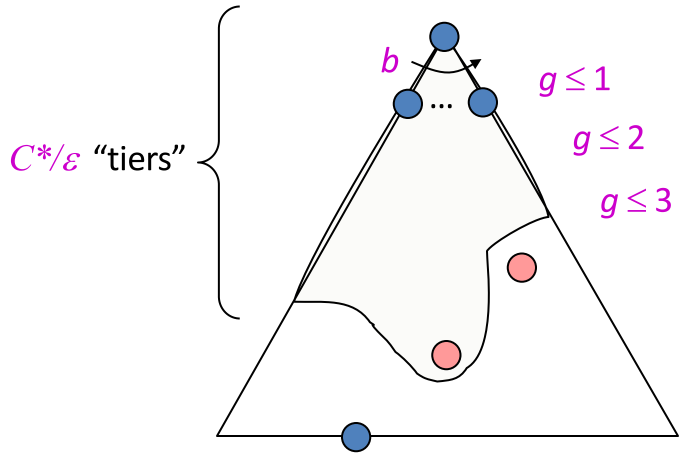
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?
  - Processes all nodes with cost less than cheapest solution!
  - If that solution costs C^* and arcs cost at least ε, then the "effective depth" is roughly C^*/ε
  - Takes time $O(b^{C^*/\varepsilon})$ (exponential in effective depth)
- How much space does the frontier take?
  - Has roughly the last tier, so $O(b^{C^*/\varepsilon})$
- Is it complete?



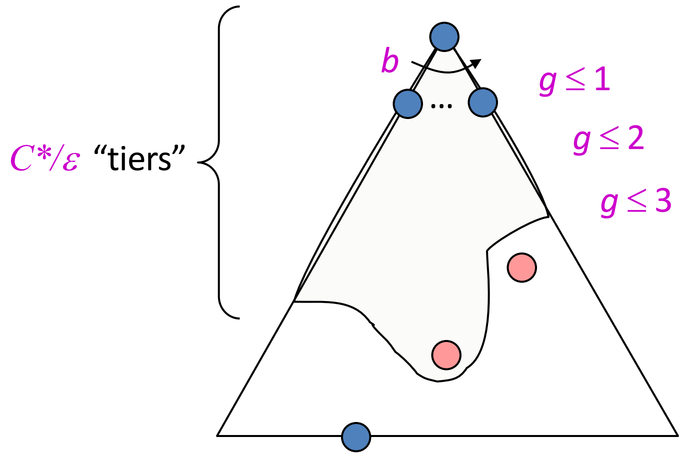
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?
  - Processes all nodes with cost less than cheapest solution!
  - If that solution costs C^* and arcs cost at least ε, then the "effective depth" is roughly C^*/ε
  - Takes time $O(b^{C^*/\varepsilon})$ (exponential in effective depth)
- How much space does the frontier take?
  - Has roughly the last tier, so $O(b^{C^*/\varepsilon})$
- Is it complete?
  - Assuming C^* is finite and $\varepsilon > 0$, yes!



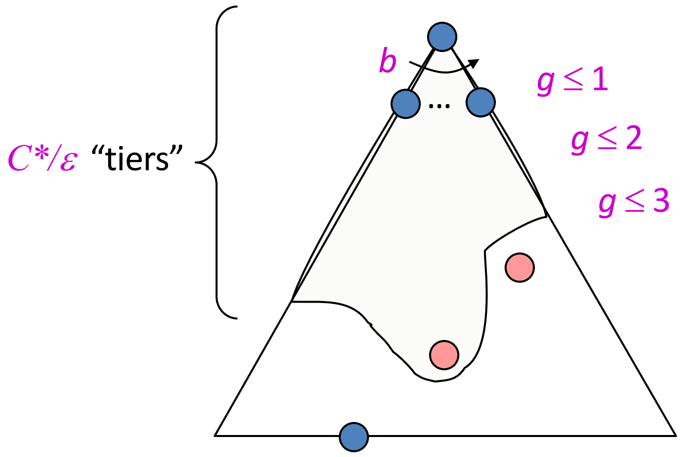
Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?
  - Processes all nodes with cost less than cheapest solution!
  - If that solution costs C^* and arcs cost at least ε, then the "effective depth" is roughly C^*/ε
  - Takes time $O(b^{C^*/\varepsilon})$ (exponential in effective depth)
- How much space does the frontier take?
  - Has roughly the last tier, so $O(b^{C^*/\varepsilon})$
- Is it complete?
  - Assuming C^* is finite and $\varepsilon > 0$, yes!
- Is it optimal?



Uniform Cost Search (UCS) Properties

- What nodes does UCS expand?
  - Processes all nodes with cost less than cheapest solution!
  - If that solution costs C^* and arcs cost at least ε, then the "effective depth" is roughly C^*/ε
  - Takes time $O(b^{C^*/\varepsilon})$ (exponential in effective depth)
- How much space does the frontier take?
  - Has roughly the last tier, so $O(b^{C^*/\varepsilon})$
- Is it complete?
  - Assuming C^* is finite and $\varepsilon > 0$, yes!
- Is it optimal?
  - Yes! (Proof next lecture via A*)



Depth-First Iterative Deepening (DFID)

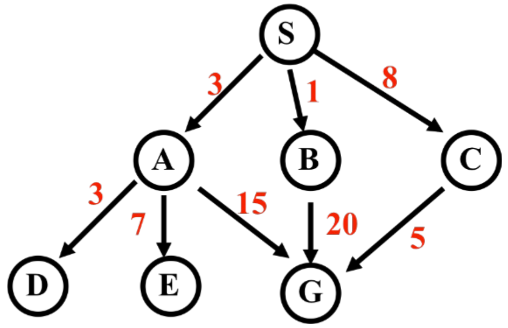
- Do DFS to depth 0, then (if no solution) DFS to depth 1, etc.
- Usually used with a tree search
- **Complete**
- **Optimal/Admissible** if all operators have unit cost, else finds shortest solution (like BFS)
- Time complexity a bit worse than BFS or DFS

  Nodes near top of search tree generated many times, but since almost all nodes are near tree bottom, worst case time complexity still exponential, O(b^d)

Depth-First Iterative Deepening (DFID)

- If branching factor is b and solution is at depth d, then nodes at depth d are generated once, nodes at depth d-1 are generated twice, etc.
  - Hence $b^d + 2b^{(d-1)} + ... + db <= b^d / (1 - 1/b)^2 = O(b^d)$.
  - If b=4, worst case is $1.78 * 4^d$, i.e., 78% more nodes searched than exist at depth d (in worst case)
- **Linear space complexity**, O(bd), like DFS
- Has advantages of BFS (completeness) and DFS (i.e., limited space, finds longer paths quickly)
- Preferred for **large state spaces** where **solution depth is unknown**

How they perform



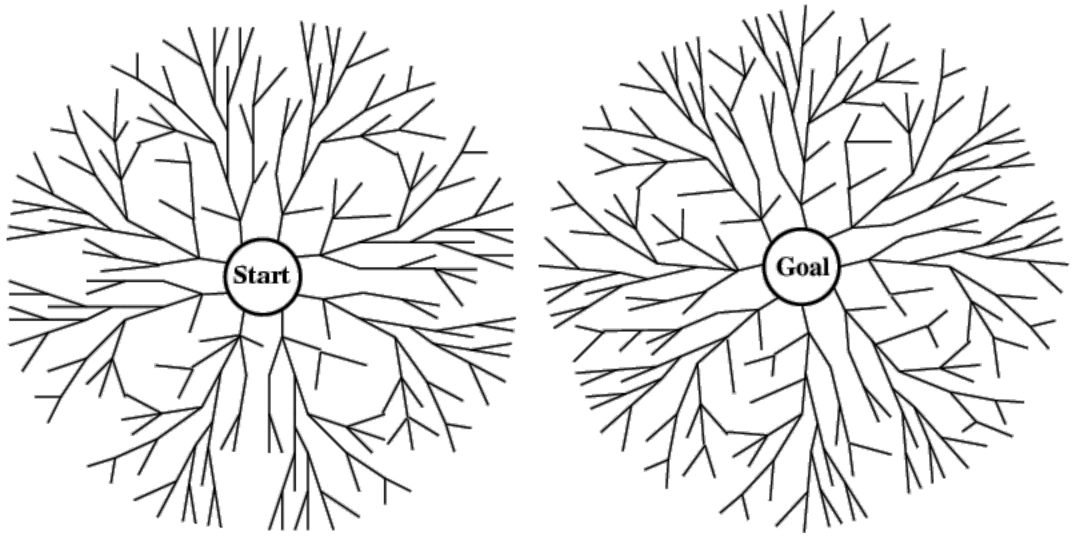
- **Depth-First Search:**
  - 4 Expanded nodes: S A D E G
  - Solution found: S A G (cost 18)
- **Breadth-First Search**:
  - 7 Expanded nodes: S A B C D E G
  - Solution found: S A G (cost 18)
- **Uniform-Cost Search**:
  - 7 Expanded nodes: S A D B C E G
  - Solution found: S C G (cost 13)

  Only uninformed search that worries about costs
- **Iterative-Deepening Search**:
  - 10 nodes expanded: S S A B C S A D E G
  - Solution found: S A G (cost 18)

Searching Backward from Goal

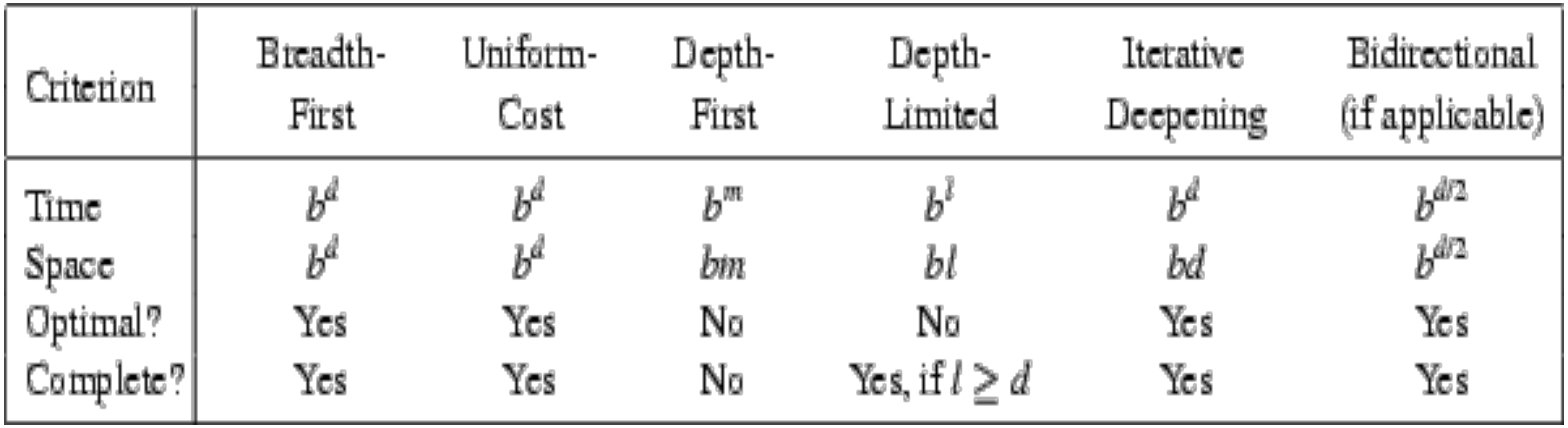
- Usually a successor function is reversible
  - i.e., can generate a node’s predecessors in graph
- If we know a single goal (rather than a goal’s properties), we could search backward to the initial state
- It might be more efficient
  - Depends on whether the graph fans in or out

Bi-directional search



- Alternate searching from the start state toward the goal and from the goal state toward the start
- Stop when the frontiers intersect
- Works well only when there are unique start & goal states
- Requires ability to generate “predecessor” states
- Can (sometimes) lead to finding a solution more quickly

Comparing Search Strategies



Criterion	Breadth-First	Uniform-Cost	Depth-First	Depth-Limited	Iterative Deepening	Bidirectional (if applicable)
Time	b^d	b^d	b^m	b^l	b^d	$b^{d/2}$
Space	b^d	b^d	bm	bl	bd	$b^{d/2}$
Optimal?	Yes	Yes	No	No	Yes	Yes
Complete?	Yes	Yes	No	Yes, if $l \geq d$	Yes	Yes

Summary

- Search in a problem space is at the heart of many AI systems
- Formalizing the search in terms of **states**, **actions**, and **goals** is key
- The simple “uninformed” algorithms we examined can be augmented to heuristics to improve them in various ways
- But for some problems, a simple algorithm is best

Informed (Heuristic) Search

- Heuristic search
- Best-first search
  - Greedy search
  - Beam search
  - A* Search
- Memory-conserving variations of A*
- Heuristic functions

Big idea: heuristic

Merriam-Webster's Online Dictionary

Heuristic (pron. \hyu-ˈris-tik\): adj. [from Greek *heuriskein* to discover] involving or serving as an aid to learning, discovery, or problem-solving by experimental and especially trial-and-error methods

The Free On-line Dictionary of Computing (15Feb98)

heuristic 1. <programming> A **rule of thumb**, simplification or educated guess that reduces or limits the search for solutions in domains that are difficult and poorly understood. Unlike algorithms, heuristics do not guarantee feasible solutions and are often used with no theoretical guarantee. 2. <algorithm> **approximation algorithm**.

From WordNet (r) 1.6

heuristic adj 1: (CS) relating to or using a heuristic rule 2: of or relating to a general formulation that serves to guide investigation [ant: algorithmic] n : a **commonsense rule** (or set of rules) intended to increase the probability of solving some problem [syn: heuristic rule, heuristic program]

Heuristics, More Formally

$h(n)$ is a **heuristic function**, that maps a state n to an estimated cost from n-to-goal

Heuristics, More Formally

$h(n)$ is a **heuristic function**, that maps a state n to an estimated cost from n-to-goal

$h(n)$ is **admissible** iff $h(n) \leq$ the lowest actual cost from n-to-goal

Heuristics, More Formally

$h(n)$ is a **heuristic function**, that maps a state n to an estimated cost from n-to-goal

$h(n)$ is **admissible** iff $h(n) \leq$ the lowest actual cost from n-to-goal

$h(n)$ is **consistent** iff

$$h(n) \leq \text{lowestcost}(n, n') + h(n')$$

Informed methods add domain-specific information

- Select best path along which to continue searching
- h(n): estimates *goodness* of node n
- h(n) = **estimated cost** (or distance) of minimal cost path from n **to a goal state**.
- Based on domain-specific information and computable from current state description that estimates how close we are to a goal

Heuristics

- **All domain knowledge** used in search is encoded in the **heuristic function, h(<node>)**
- Examples:
  - 8-puzzle: number of tiles out of place
  - 8-puzzle: sum of distances each tile is from its goal
  - Missionaries & Cannibals: # people on starting river bank
- In general
  - $h(n) \geq 0$ for all nodes n
  - $h(n)$ = 0 implies that n is a goal node
  - $h(n) = \infty$ implies n is a dead-end that can't lead to goal

Heuristics for 8-puzzle

Misplaced Tiles Heuristic

(not including the blank)

Current State

3	2	8
4	5	6
7	1	

Goal State

1	2	3
4	5	6
7	8	

3	2	8
4	5	6
7	1	

3 tiles are not where they need to be

- Three tiles are misplaced (the 3, 8, and 1) so heuristic function evaluates to 3
- Heuristic says that it *thinks* a solution may be available in **3 or more** moves
- Very rough estimate, but easy to calculate

h = 3

Heuristics for 8-puzzle

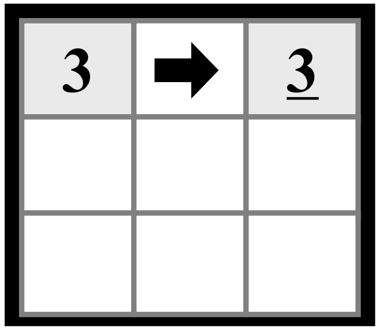
Manhattan Distance (not including the blank)

Current State

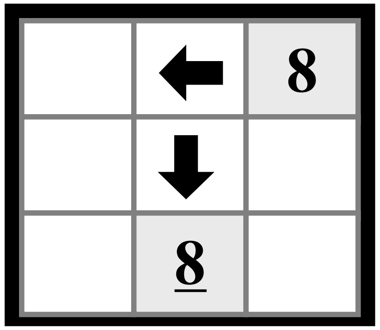
3	2	8
4	5	6
7	1	

Goal State

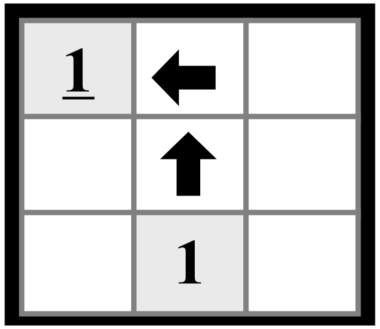
1	2	3
4	5	6
7	8	



2 spaces



3 spaces



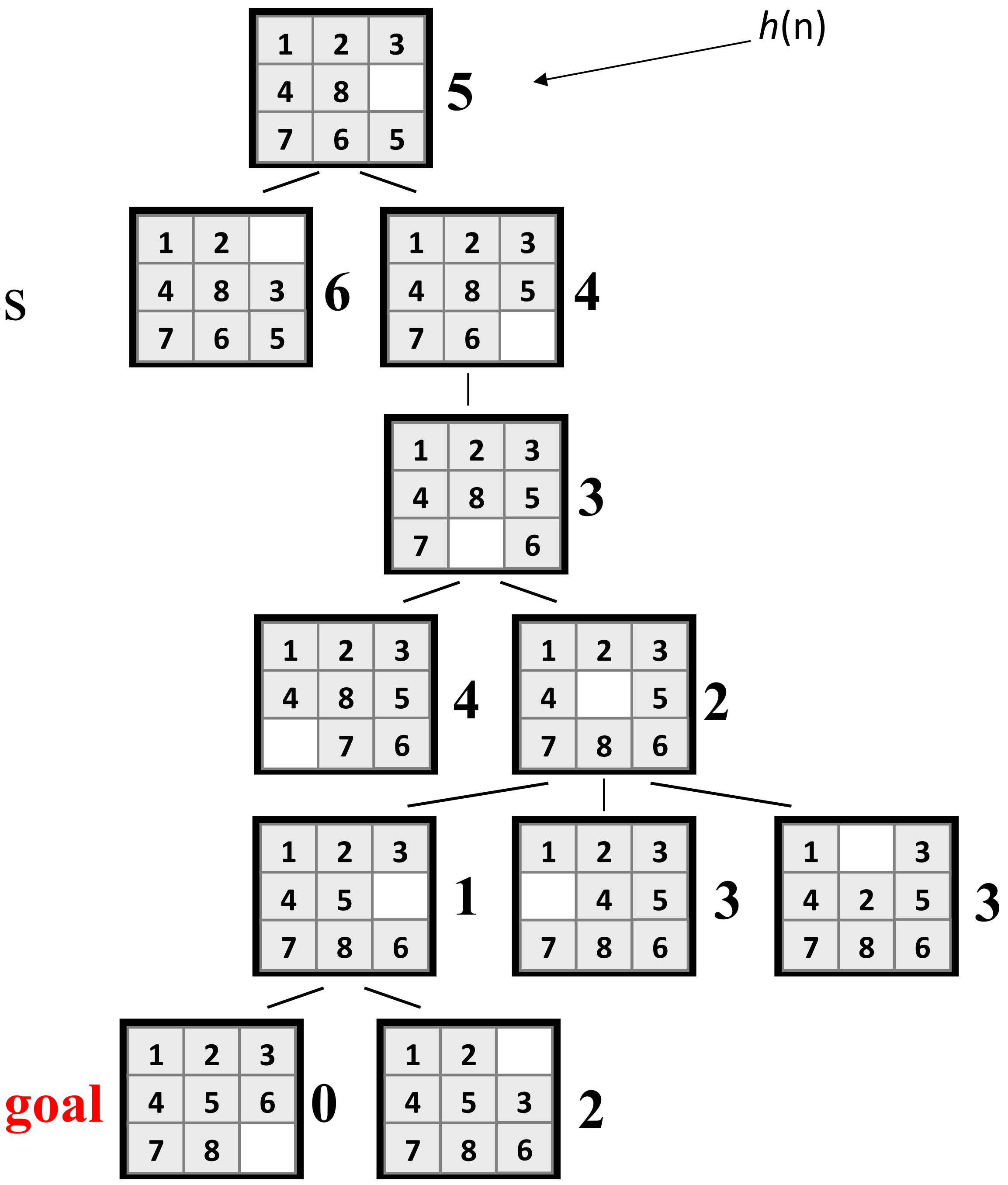
3 spaces

Total 8

- The **3**, **8** and **1** tiles are misplaced (by 2, 3, and 3 steps) so the heuristic function evaluates to 8
- Heuristic says that it *thinks* a solution may be available in just 8 more moves.
- The misplaced heuristic's value is 3

We can use heuristics to guide search

Manhattan Distance heuristic helps us quickly find a solution to the 8-puzzle



Best-first search

- Search algorithm that improves **depth-first search** by expanding most promising node chosen according to heuristic rule

- Order nodes on Fringe list by increasing value of an evaluation function, **f(n)**, incorporating domain-specific information

Best-first search

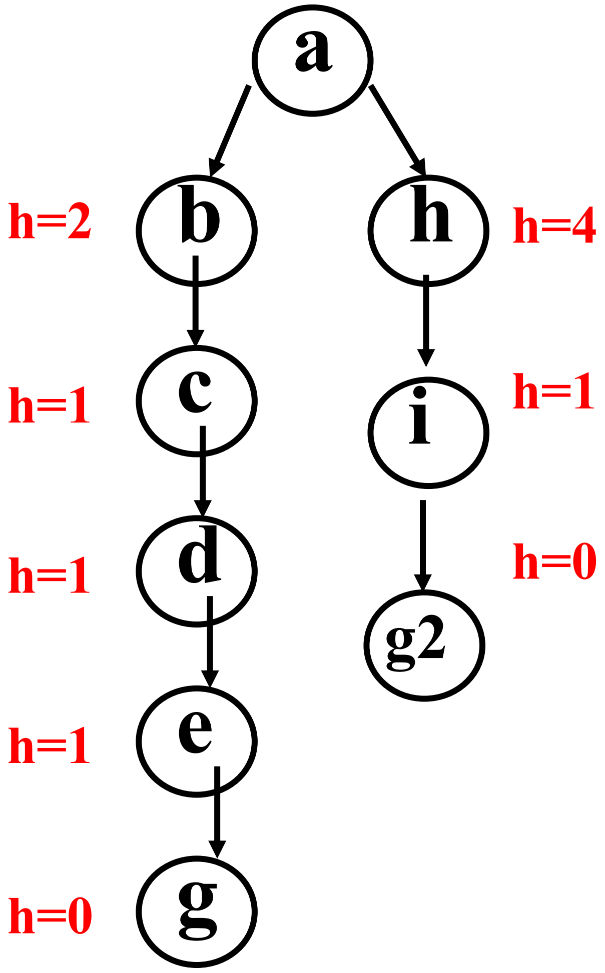
- Search algorithm that improves **depth-first search** by expanding most promising node chosen according to heuristic rule
- Order nodes on Fringe list by increasing value of an evaluation function, **f(n)**, incorporating domain-specific information
- This is a generic way of referring to the class of informed methods

Greedy best first search

- A greedy algorithm makes locally optimal choices in hope of finding a global optimum
- Uses evaluation function *f(n) = h(n)*, sorting nodes by increasing values of *f*
- Selects node to expand appearing **closest** to goal (i.e., node with smallest f value)
- Not complete as can end up in dead end when solution not found
- Not admissible, as in example
  - Assume arc costs = 1, greedy search finds goal g, with solution cost of 5
  - Optimal solution is path to goal with cost 3

Greedy best first search example

- Proof of non-admissibility
  - Assume arc costs = 1, greedy search finds goal g, with solution cost of 5
  - Optimal solution is path to goal with cost 3



Greedy best first search example

- Makes locally optimal choices at each step based on the current information and do not reconsider past decisions.
- Once a greedy algorithm makes a choice and moves to the next step, it does not go back to reconsider or explore alternative paths. In some cases, **they can get stuck in local optima or suboptimal solutions.**
- If fails to find a path to the goal, then the chosen path based on the heuristic did not lead to a solution. In such cases, the algorithm may terminate without finding a solution or may need to be modified to explore alternative paths, possibly incorporating backtracking, to improve its search capabilities.

Beam search

- Instead of picking one child per iteration, it expands k number of children, **in parallel.**
- Use evaluation function f(n), but maximum size of the nodes list is k, a fixed constant
- Only keep k best nodes as candidates for expansion, discard rest
- k is the *beam width*
- More space efficient than greedy search, but may discard nodes on a solution path
- As k increases, approaches best first search
- Complete?
- Admissible?

Beam search

- Instead of picking one child per iteration, it expands k number of children**, in parallel.**
- Use evaluation function f(n), but maximum size of the nodes list is k, a fixed constant
- Only keep k best nodes as candidates for expansion, discard rest
- k is the *beam width*
- More space efficient than greedy search, but may discard nodes on a solution path
- As k increases, approaches best first search
- Not complete
- Not admissible

We’ve *got* to be able to do better, right?

Let’s think about car trips…

A* Search

Use an evaluation function

$$f(n) = g(n) + h(n)$$

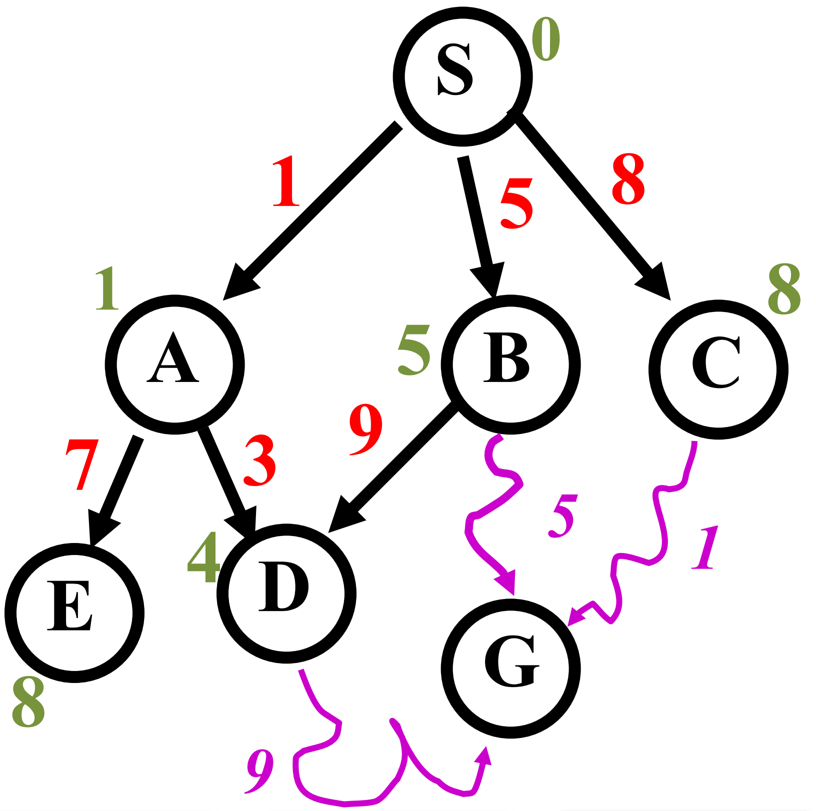
estimated total cost from start to goal via state n = minimal-cost path from the start state to state n + cost estimate from state n to the goal

A* Search

- Use as an evaluation function

  f(n) = g(n) + h(n)

- g(n) = minimal-cost path from the start state to state n
- g(n) adds "breadth-first" term to evaluation function
- Ranks nodes on search frontier by estimated cost of solution from start node ***via given node*** to goal



g(d)=4
h(d)=9
f(d)=13

g(b)=5
h(b)=5
f(d)=10

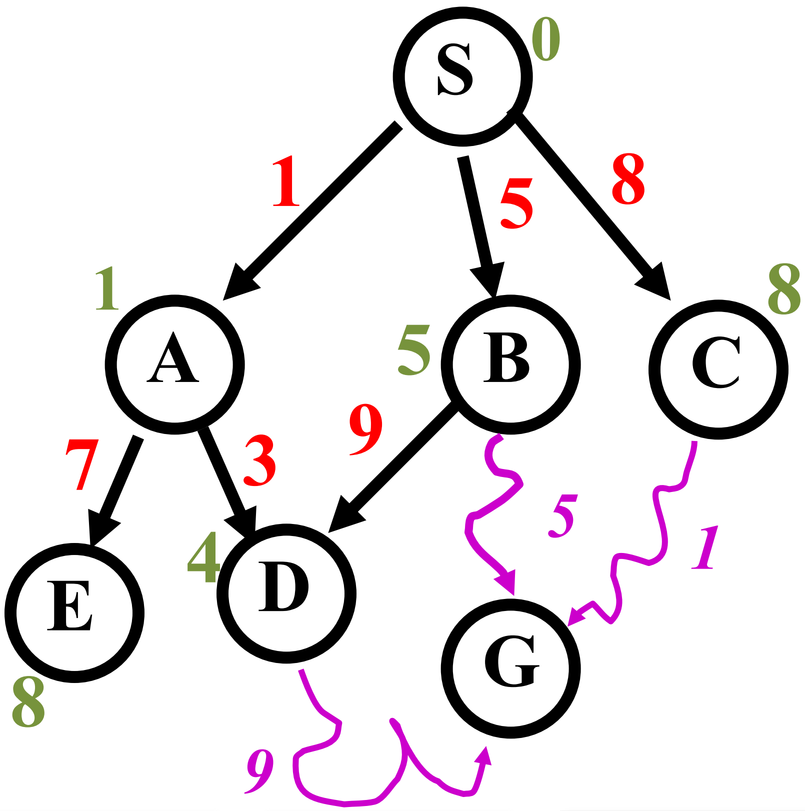
g(c)=8
h(c)=1
f(c)=9

A* Search

- Use as an evaluation function

  f(n) = g(n) + h(n)

- g(n) = minimal-cost path from the start state to state n
- g(n) adds "breadth-first" term to evaluation function
- Ranks nodes on search frontier by estimated cost of solution from start node ***via given node*** to goal



g(d)=4	*g(b)=5*	*g(c)=8*
h(d)=9	*h(b)=5*	*h(c)=1*
f(d)=13	*f(d)=10*	*f(c)=9*

C is chosen next to expand

A* Search

- Use an evaluation function

$$\mathbf{f(n) = g(n) + h(n)}$$

estimated total cost from start to goal via state n = minimal-cost path from the start state to state n + cost estimate from state n to the goal

- g(n) term adds "breadth-first" component to evaluation function
- Ranks nodes on search frontier by estimated cost of solution from start node *via given node* to goal
- Not complete if h(n) can = ∞
- Is it admissible?

A*

- Pronounced *"a star"*
- h is **admissible** when h(n) <= h*(n) holds
  - **h*(n)** = *true cost* of *minimal cost path* from n to a goal
- Using an admissible heuristic guarantees that 1st solution found will be an **optimal** one
  - With an admissible heuristic, A* is cost-optimal
- A* is **complete** whenever branching factor is finite and every action has fixed, positive cost
- A* is **admissible**

Hart, P. E.; Nilsson, N. J.; Raphael, B. (1968). "A Formal Basis for the Heuristic Determination of Minimum Cost Paths". *IEEE Transactions on Systems Science and Cybernetics SSC4* **4** (2): 100–107.

Implementing A*

Q: Can this be an instance of our general search algorithm?

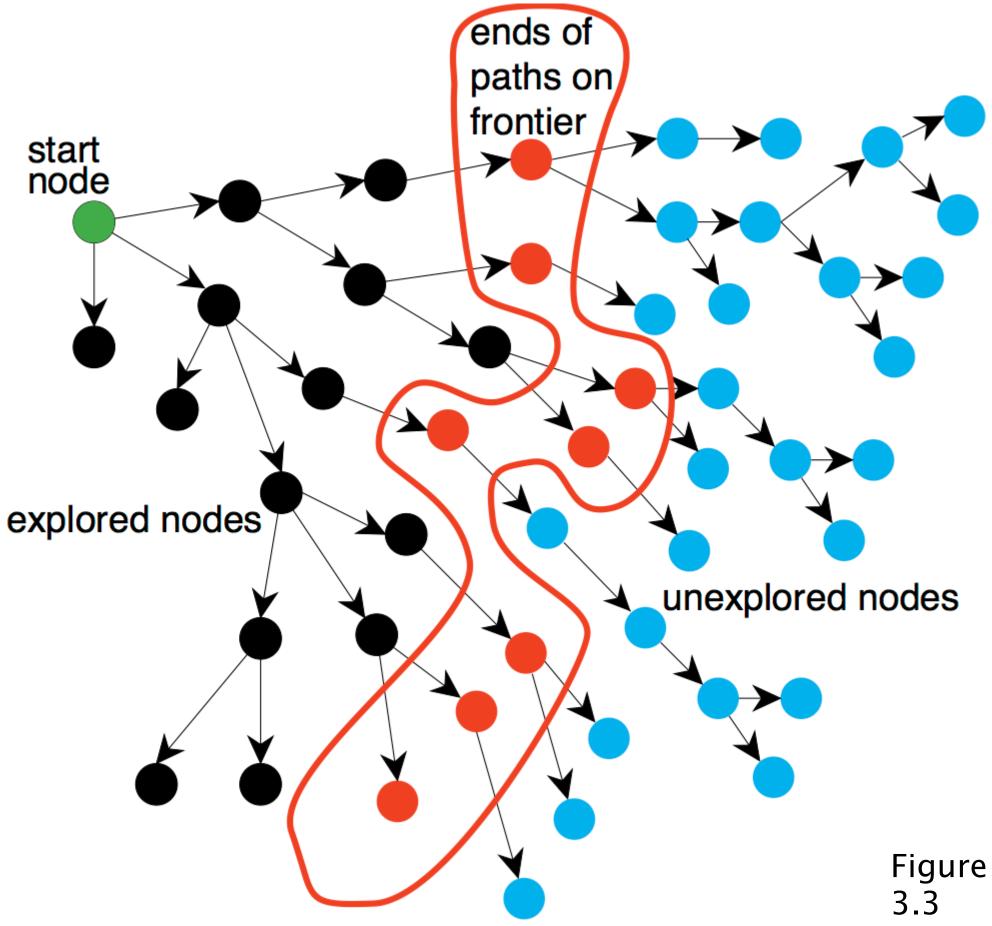


Figure 3.3

Implementing A*

Q: Can this be an instance of our general search algorithm?

A: Yup! Just make the fringe a priority queue ordered by $f(n)$

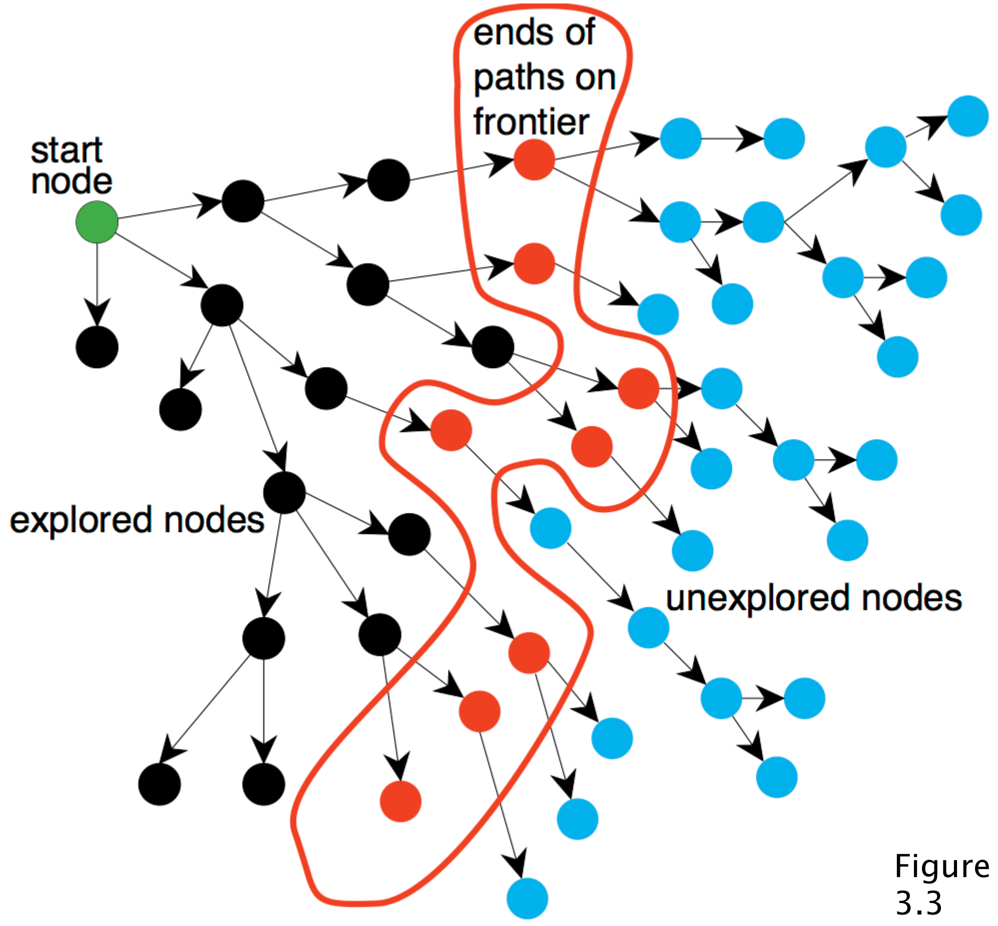


Figure 3.3

Alternative A* Pseudo-code

```
1 Put the start node S on the nodes list, called OPEN
2 If OPEN is empty, exit with failure
3 Select node in OPEN with minimal f(n) and place on CLOSED
4 If n is a goal node, collect path back to start and stop
5 Expand n, generating all its successors and attach to them
  pointers back to n.  For each successor n' of n
   1 If n’ not already on OPEN or CLOSED
       • put n' on OPEN
       • compute h(n'),  g(n')=g(n)+ c(n, n'),  f(n')=g(n')+h(n')
   2 If n’ already on OPEN or CLOSED and if g(n') is lower for new
     version of n', then:
       • Redirect pointers backward from n’ on path with lower g(n’)
       • Put n' on OPEN
```

Next class

IS A HEURISTIC ADMISSIBLE?