



香港科技大学(广州)
THE HONG KONG
UNIVERSITY OF SCIENCE AND
TECHNOLOGY (GUANGZHOU)

Design and Analysis of Algorithms

Jing Tang | DSAA 2043 Fall 2024

Shortest Paths

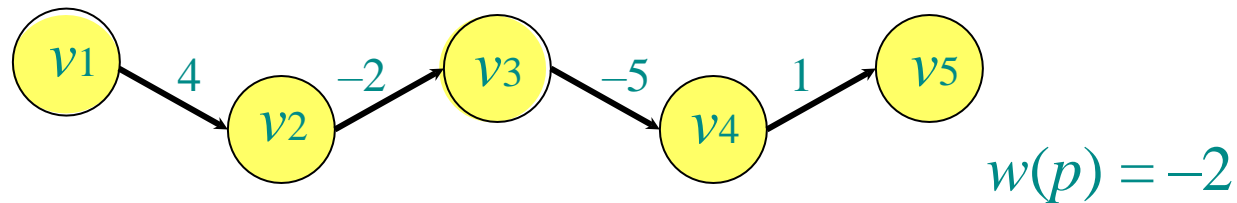
- Define the **single source shortest path** problem.
- Optimal substructure, Triangle inequality and **Relaxation**.
- The Dijkstra algorithm and its extension.

Consider a digraph $G = (V, E)$ with edge-weight function $w : E \rightarrow \mathbb{R}$.

The **weight** of path $p = v_1 \rightarrow v_2 \rightarrow \dots \rightarrow v_k$ is defined to be

$$w(p) = \sum_{i=1}^{k-1} w(v_i, v_{i+1}).$$

Example:



A **shortest path** from u to v is a path of minimum weight from u to v .

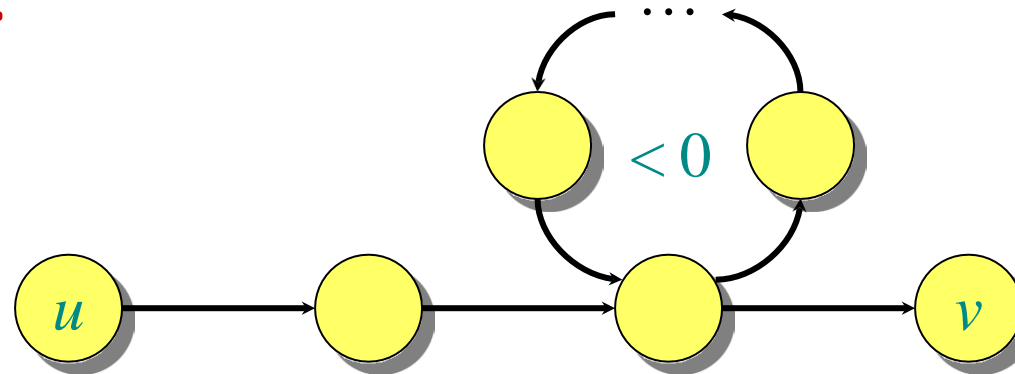
The **shortest-path weight** from u to v is defined as:

$$\delta(u, v) = \min\{w(p) : p \text{ is a path from } u \text{ to } v\}.$$

Note: $\delta(u, v) = \infty$ if no path from u to v exists.

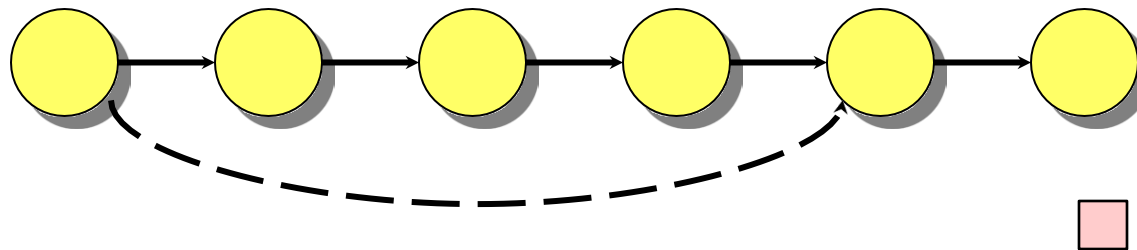
If a graph G contains a negative-weight cycle, then some shortest paths do not exist.

Example:



Theorem. A subpath of a shortest path is a shortest path.

Proof. Cut and paste:



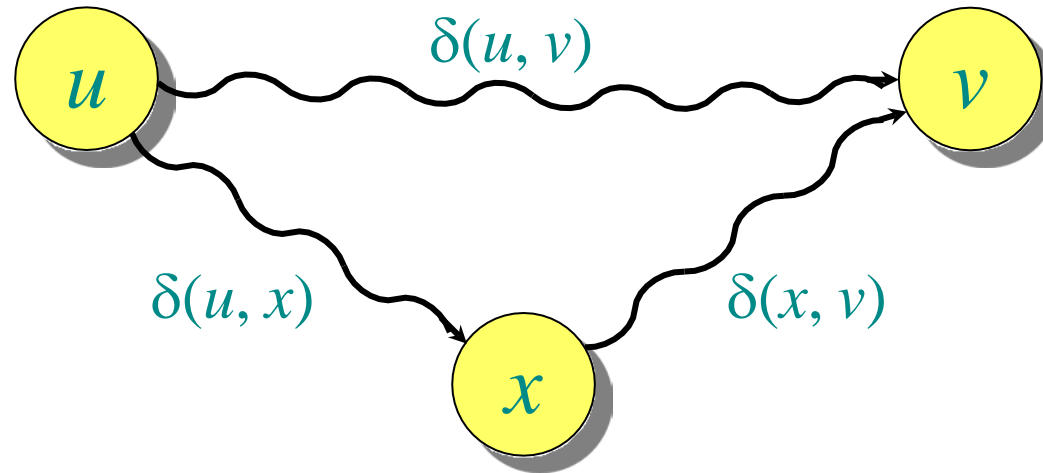
If v_j on optimal path from v_0 to v_n : $\delta(v_0, v_n) = \delta(v_0, v_j) + \delta(v_j, v_n)$.

If the sub-path v_i to v_j is not optimal, then by finding a shorter path from v_i to v_j we can strictly improve the original path.

Theorem. For all $u, v, x \in V$, we have

$$\delta(u, v) \leq \delta(u, x) + \delta(x, v).$$

Proof.



If u not on shortest path from s to t : $\delta(s, t) < \delta(s, u) + \delta(u, t)$.

u is on shortest path from s to t iff $\delta(s, t) = \delta(s, u) + \delta(u, t)$.

Problem. Assume that $w(u, v) \geq 0$ for all $(u, v) \in E$. (Hence, all shortest-path weights must exist.) From a given source vertex $s \in V$, find the shortest-path weights $\delta(s, v)$ for all $v \in V$.

IDEA: Greedy.

1. Maintain a set S of vertices whose shortest-path distances from s are known.
2. At each step, add to S the vertex $v \in V - S$ whose distance estimate from s is minimum.
3. Update the distance estimates of vertices adjacent to v .

$d[s] \leftarrow 0$

for each $v \in V - \{s\}$

do $d[v] \leftarrow \infty$

$S \leftarrow \emptyset$

$Q \leftarrow V$

▷ Q is a priority queue maintaining $V - S$,
keyed on $d[v]$

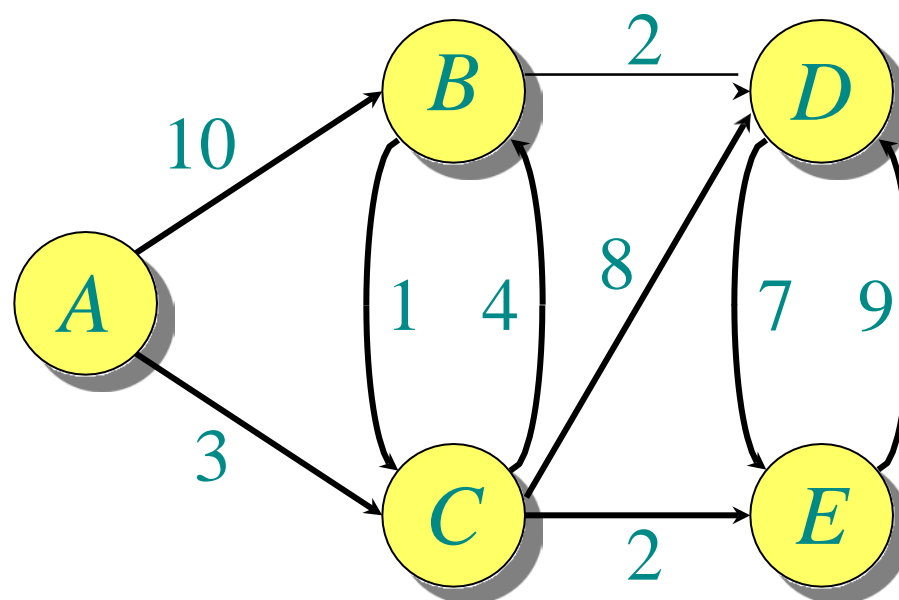
```
 $d[s] \leftarrow 0$   
for each  $v \in V - \{s\}$   
  do  $d[v] \leftarrow \infty$   
 $S \leftarrow \emptyset$   
 $Q \leftarrow V$  ▷  $Q$  is a priority queue maintaining  $V - S$ ,  
  keyed on  $d[v]$   
while  $Q \neq \emptyset$   
  do  $u \leftarrow \text{EXTRACT-MIN}(Q)$   
     $S \leftarrow S \cup \{u\}$   
    for each  $v \in \text{Adj}[u]$   
      do if  $d[v] > d[u] + w(u, v)$   
        then  $d[v] \leftarrow d[u] + w(u, v)$ 
```

```
 $d[s] \leftarrow 0$ 
for each  $v \in V - \{s\}$ 
do  $d[v] \leftarrow \infty$ 
 $S \leftarrow \emptyset$ 
 $Q \leftarrow V$        $\triangleright Q$  is a priority queue maintaining  $V - S$ ,
                     keyed on  $d[v]$ 
while  $Q \neq \emptyset$ 
do  $u \leftarrow \text{EXTRACT-MIN}(Q)$ 
   $S \leftarrow S \cup \{u\}$ 
  for each  $v \in \text{Adj}[u]$ 
  do if  $d[v] > d[u] + w(u, v)$ 
     then  $d[v] \leftarrow d[u] + w(u, v)$       relaxation step
```

\nwarrow Implicit DECREASE-KEY

Example of Dijkstra's algorithm

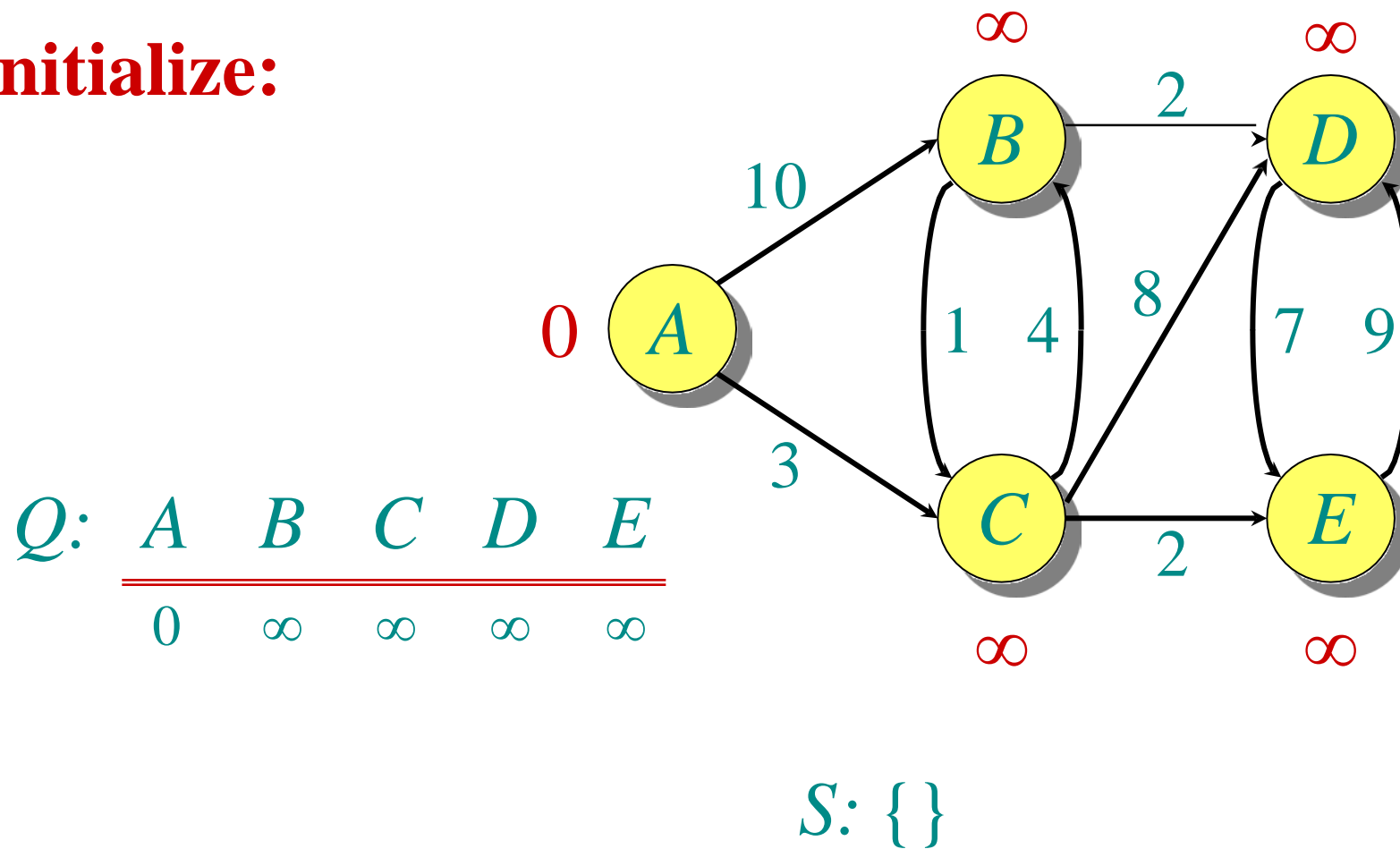
**Dijkstra can only handle
graphs with nonnegative
edge weights:**



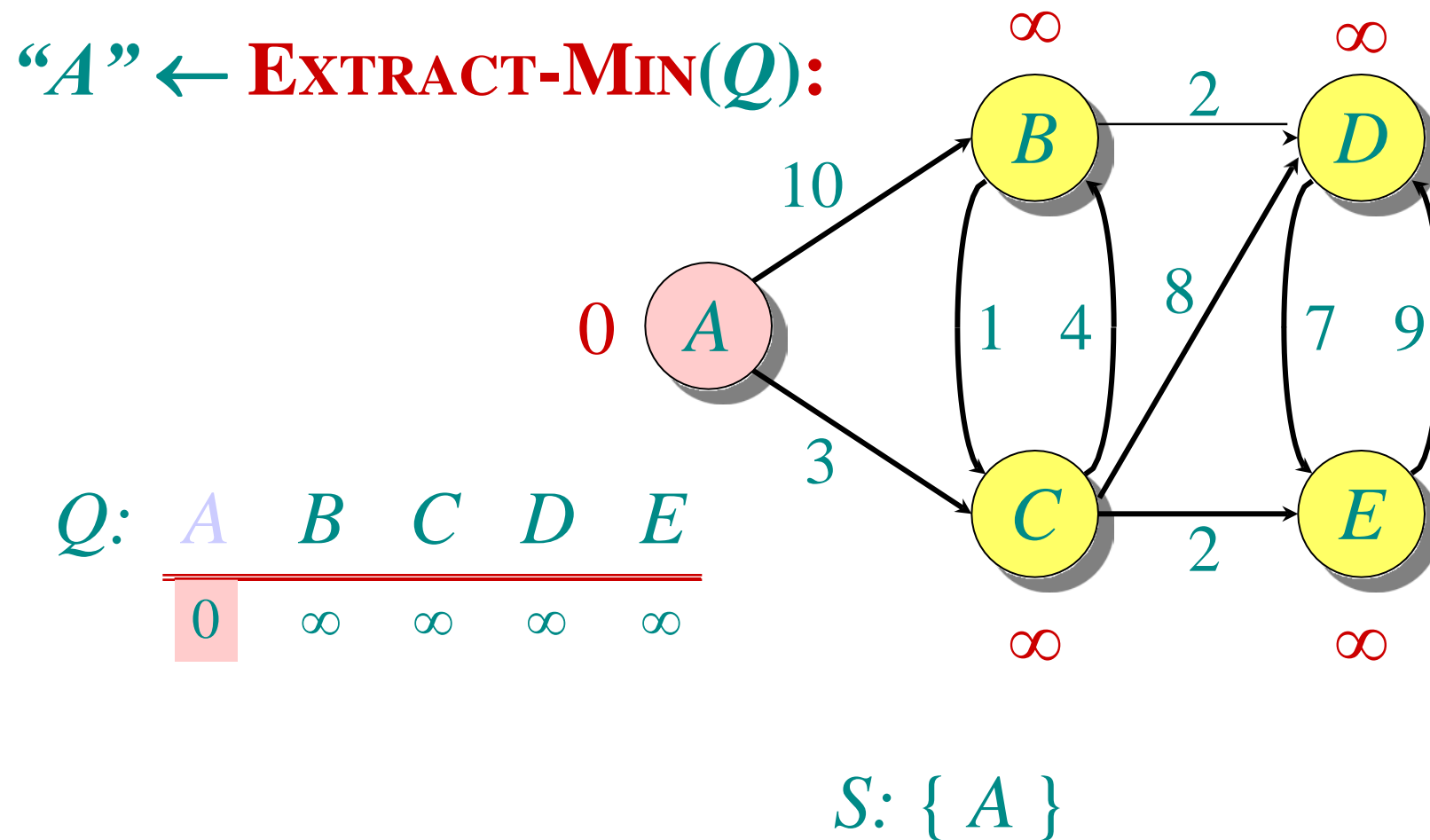
Try to think why?

Example of Dijkstra's algorithm

Initialize:

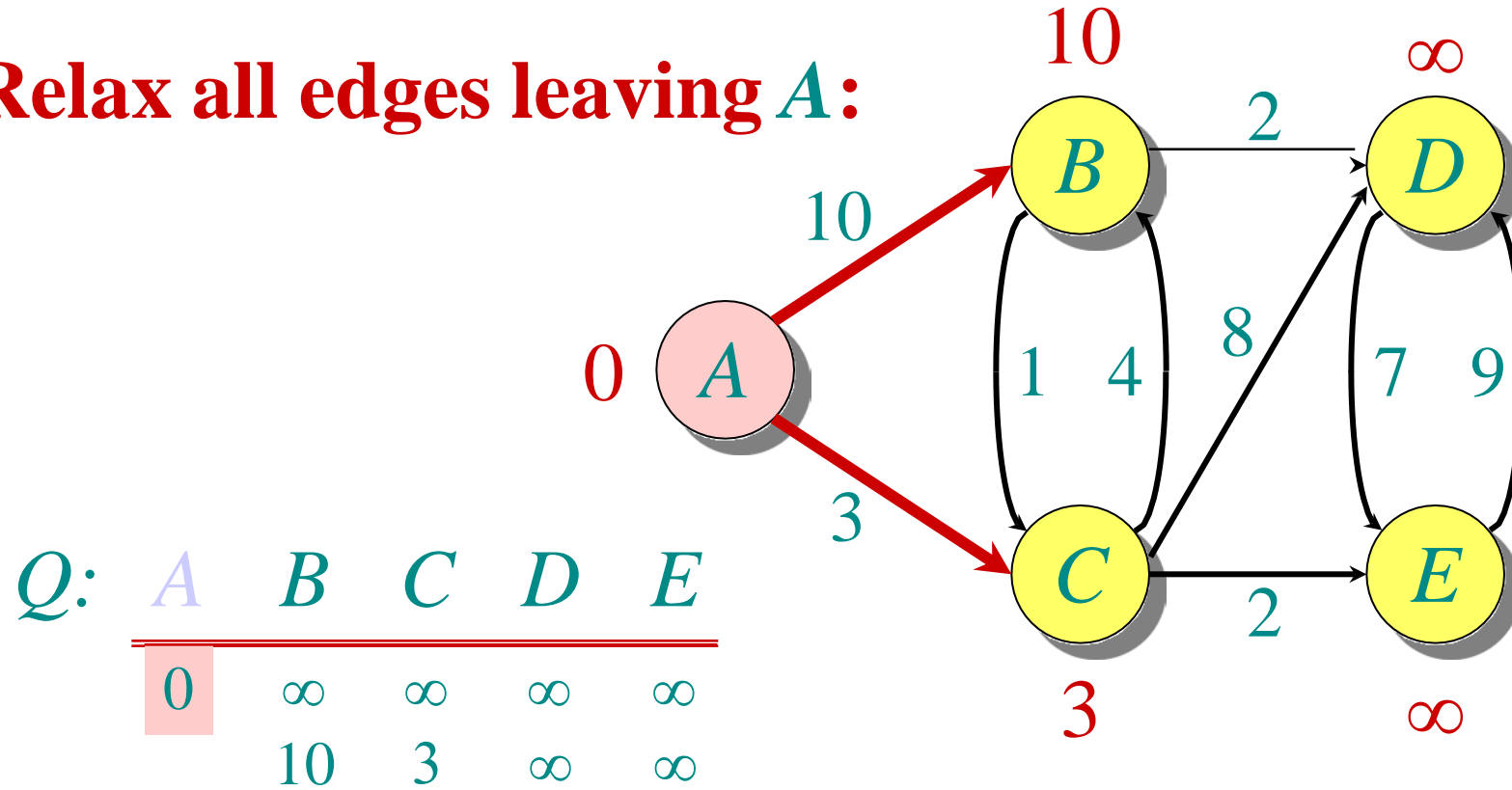


Example of Dijkstra's algorithm



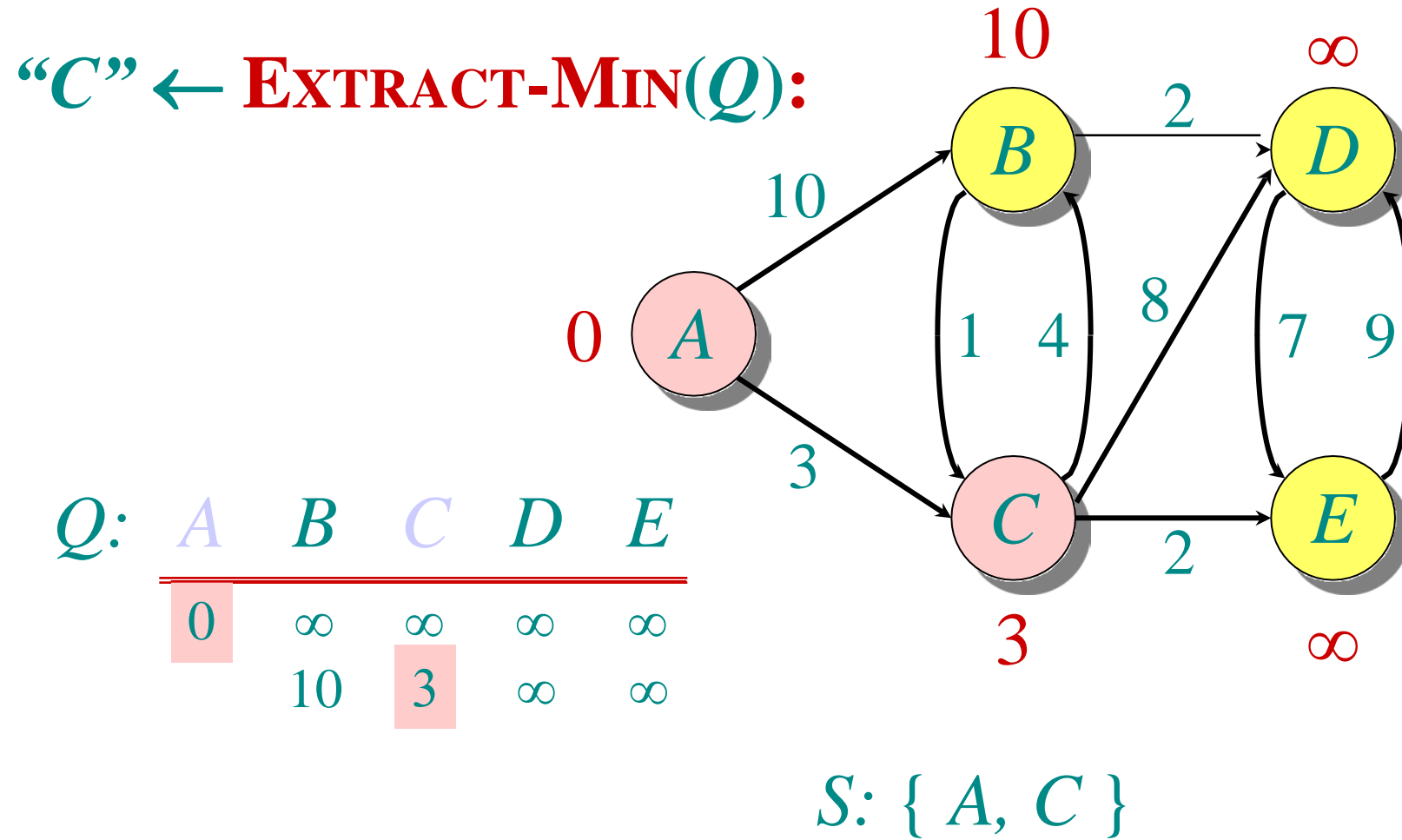
Example of Dijkstra's algorithm

Relax all edges leaving **A**:



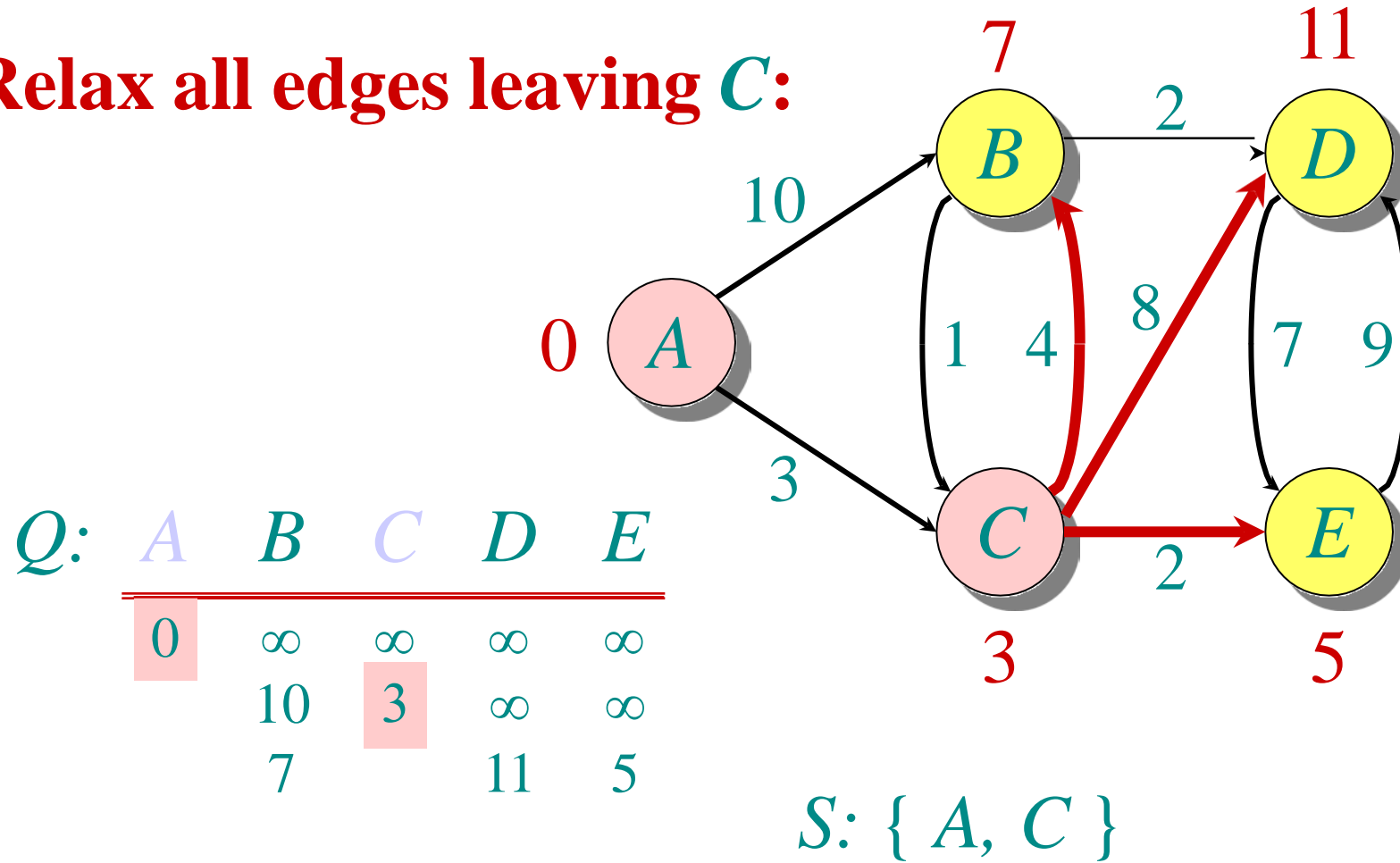
S: { **A** }

Example of Dijkstra's algorithm

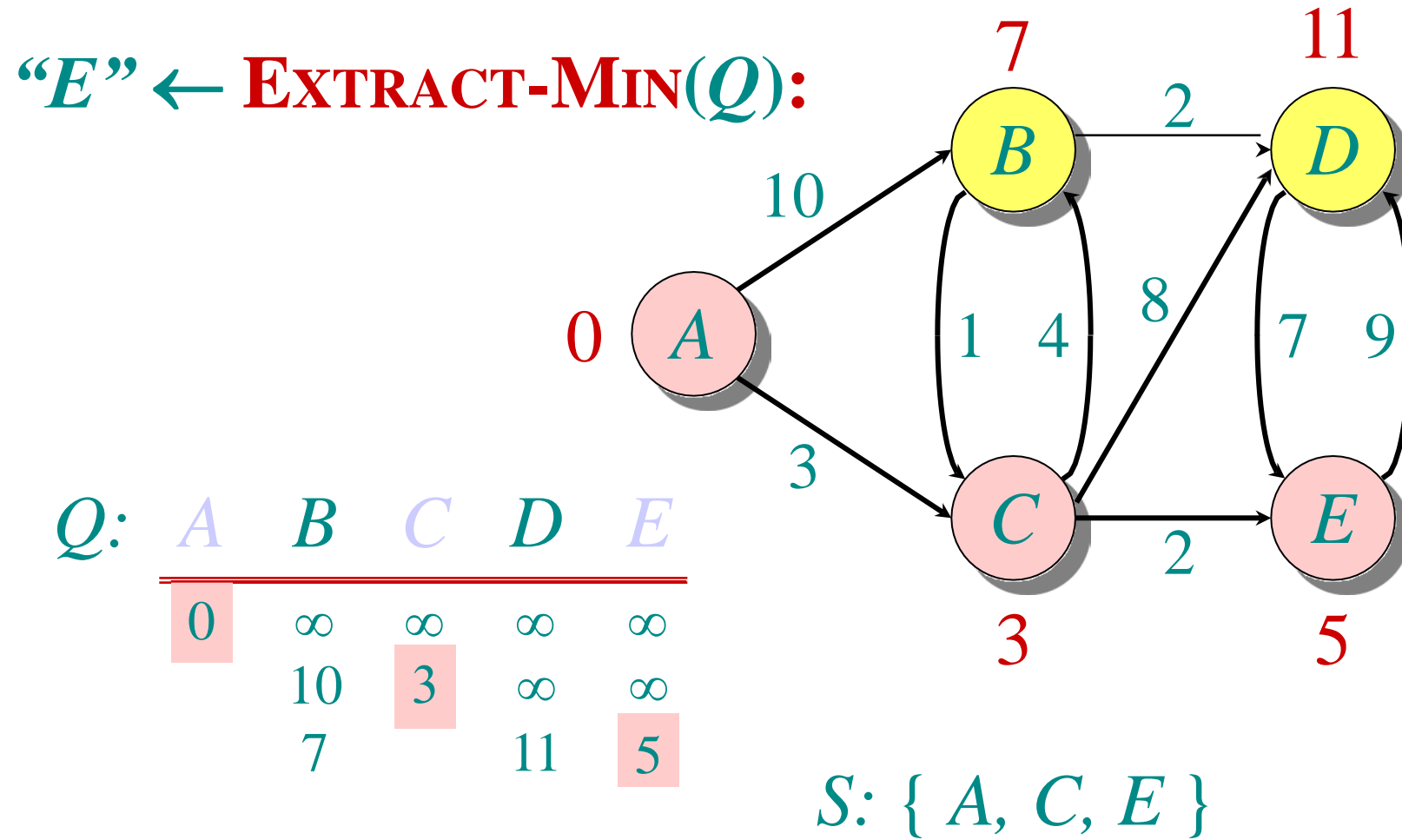


Example of Dijkstra's algorithm

Relax all edges leaving **C**:

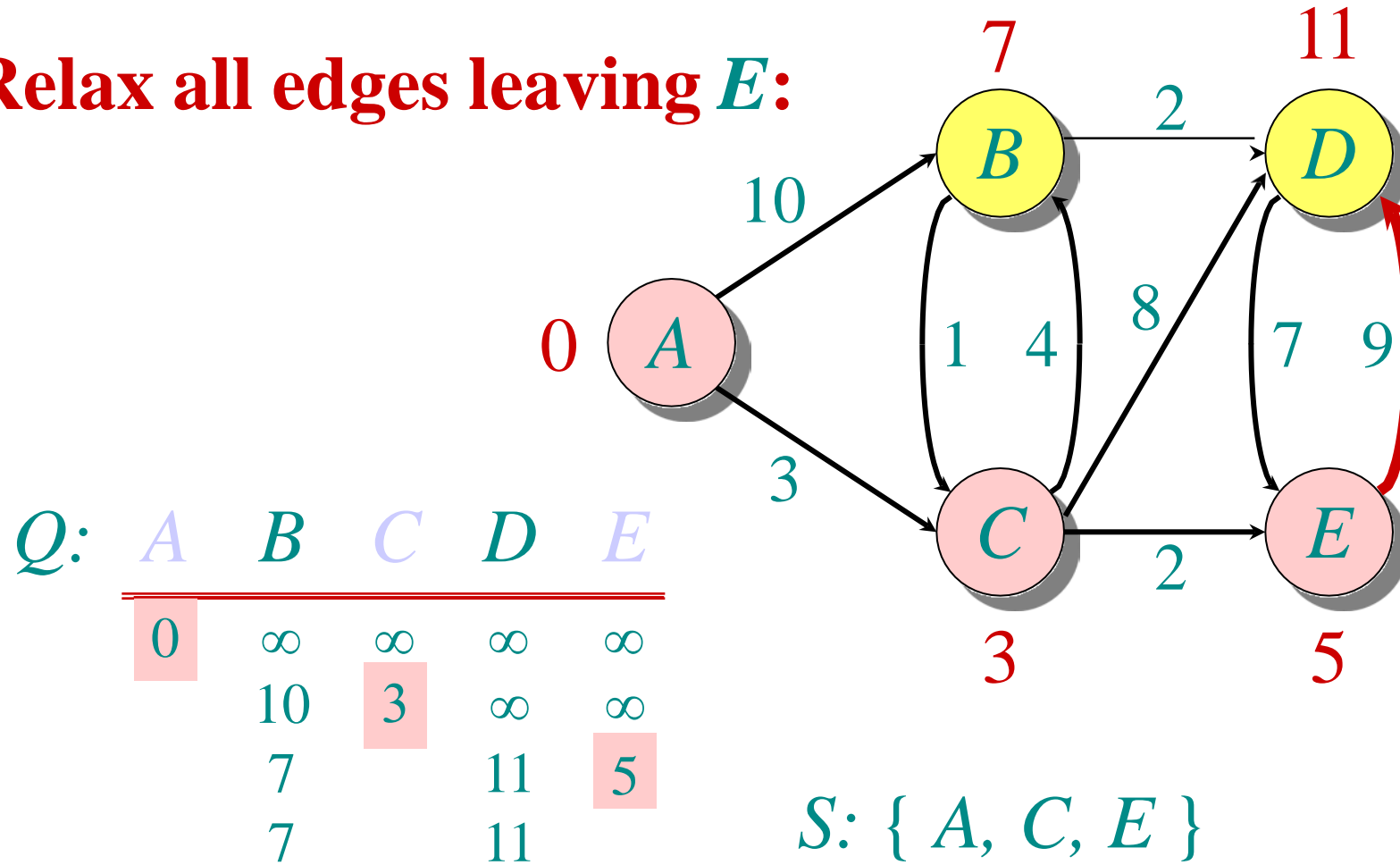


Example of Dijkstra's algorithm

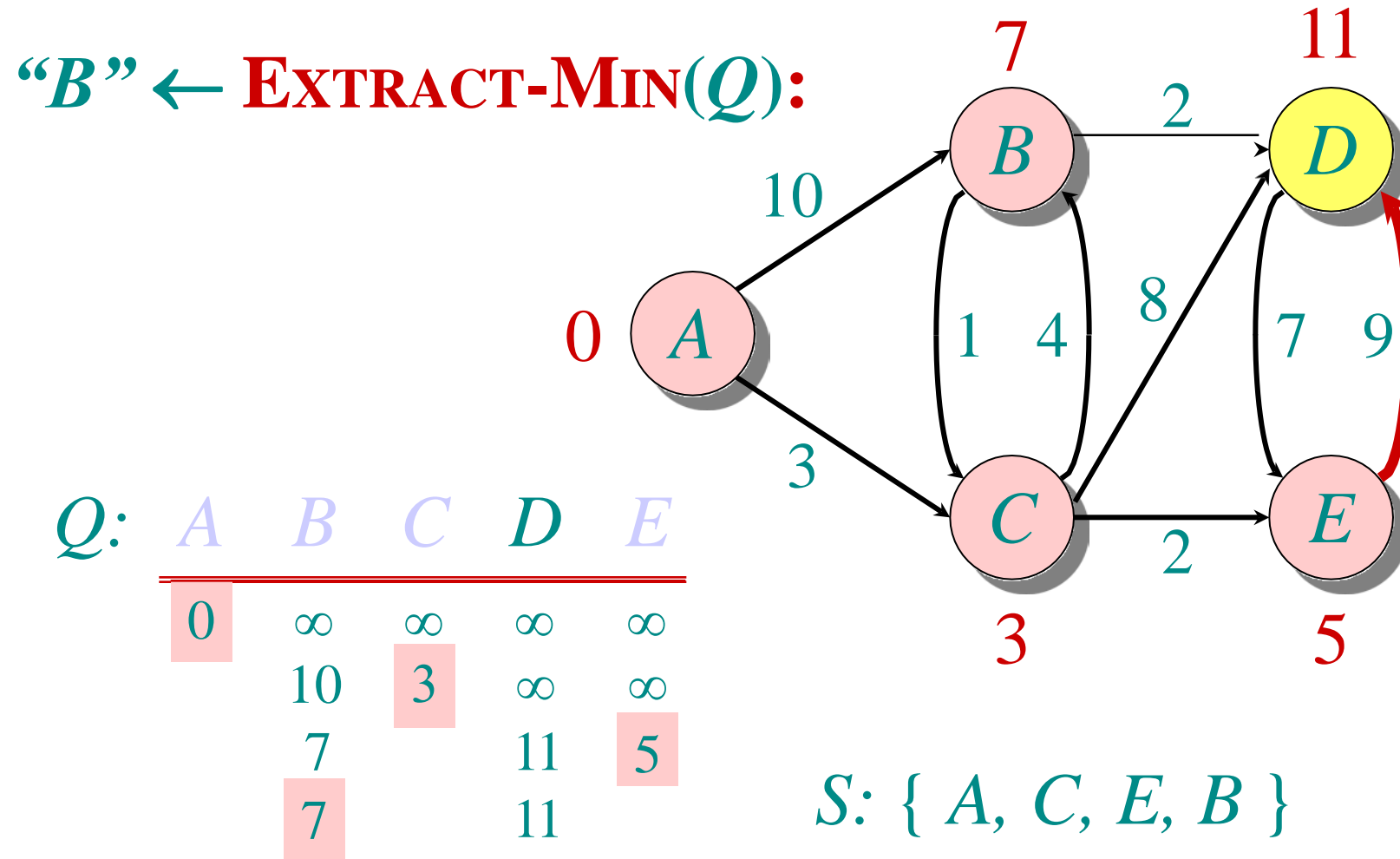


Example of Dijkstra's algorithm

Relax all edges leaving *E*:

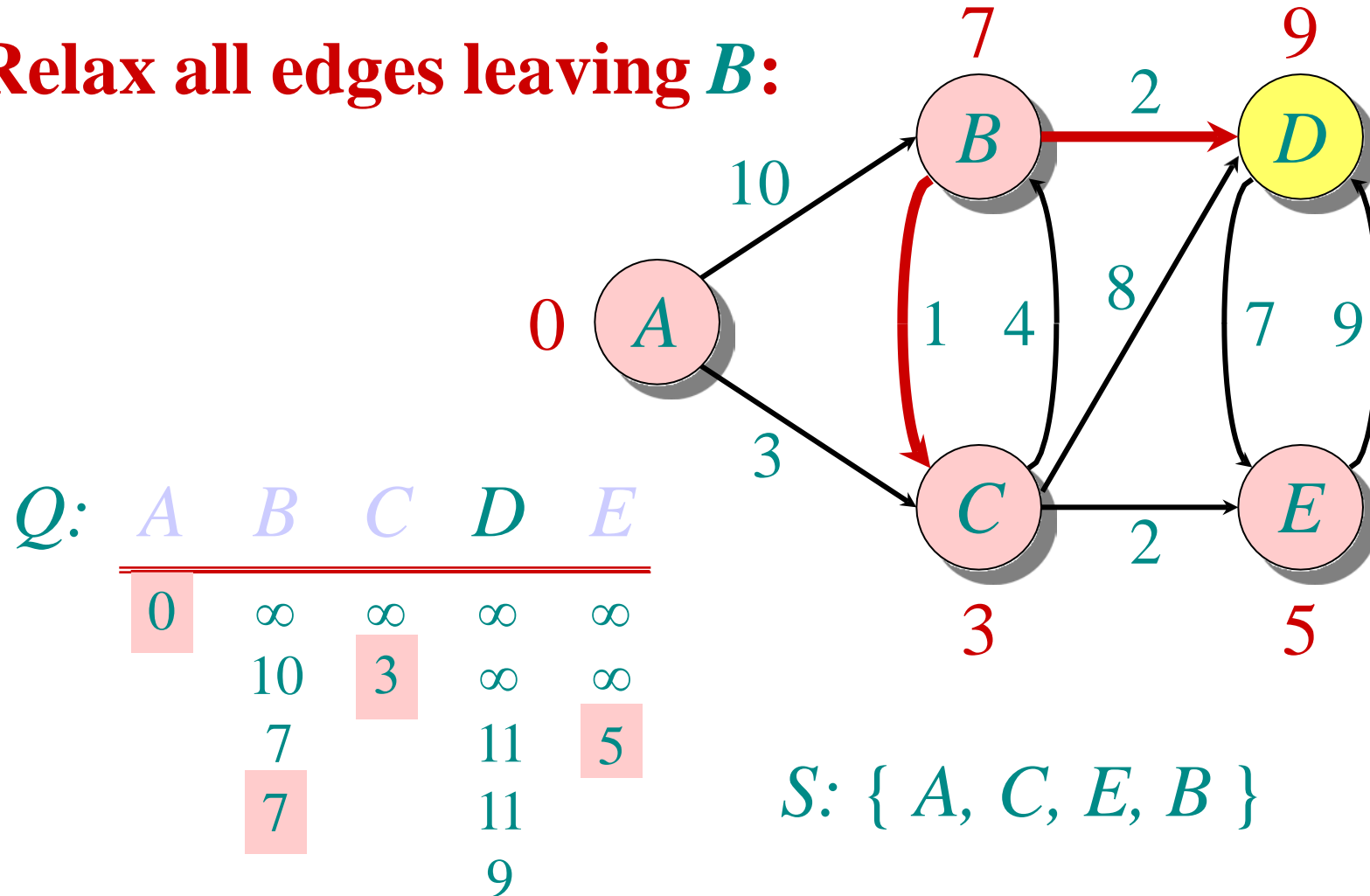


Example of Dijkstra's algorithm



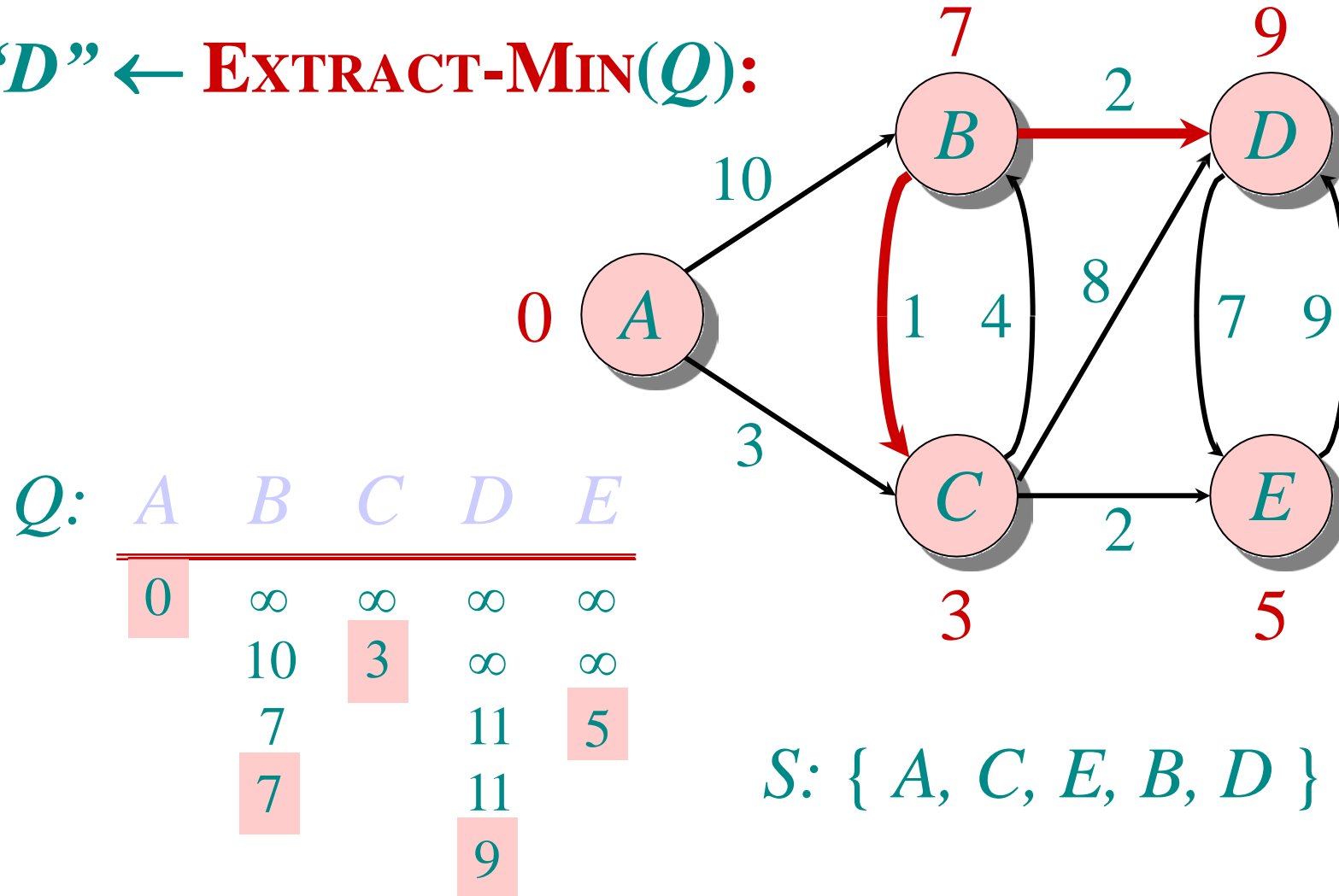
Example of Dijkstra's algorithm

Relax all edges leaving *B*:



Example of Dijkstra's algorithm

“D” \leftarrow **EXTRACT-MIN(Q):**



Lemma. Initializing $d[s] \leftarrow 0$ and $d[v] \leftarrow \infty$ for all $v \in V - \{s\}$ establishes $d[v] \geq \delta(s, v)$ for all $v \in V$, and this invariant is maintained over any sequence of relaxation steps.


Proof. Suppose not. Let v be the **first** vertex for which $d[v] < \delta(s, v)$, and let u be the vertex that caused $d[v]$ to change: $d[v] = d[u] + w(u, v)$. Then,

$d[v] < \delta(s, v)$	supposition
$\leq \delta(s, u) + \delta(u, v)$	triangle inequality
$\leq \delta(s, u) + w(u, v)$	sh. path \leq specific path
$\leq d[u] + w(u, v)$	v is first violation

Contradiction.

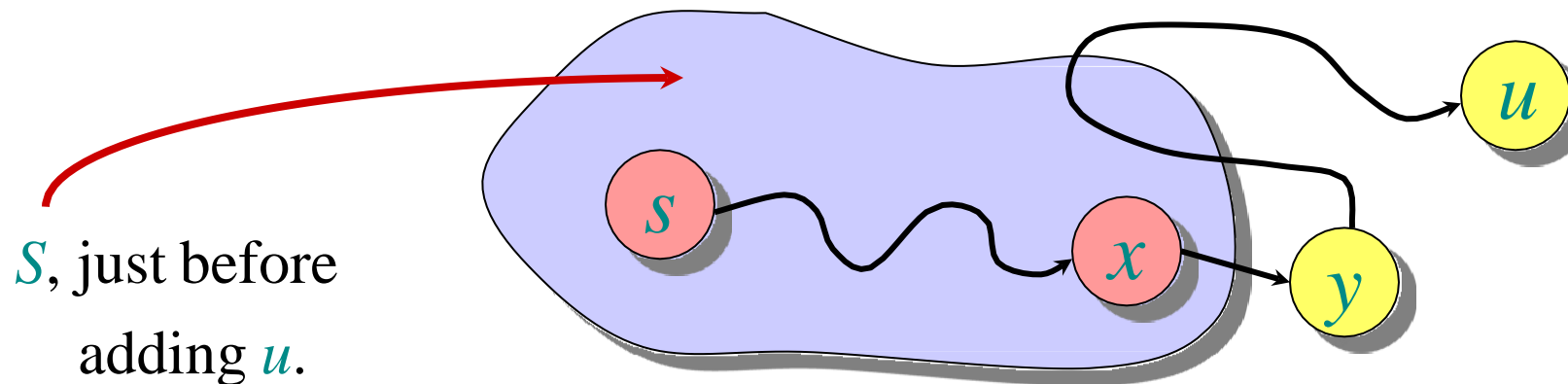
Lemma. Let u be v 's predecessor on a shortest path from s to v . Then, if $d[u] = \delta(s, u)$ and edge (u, v) is relaxed, we have $d[v] = \delta(s, v)$ after the relaxation.

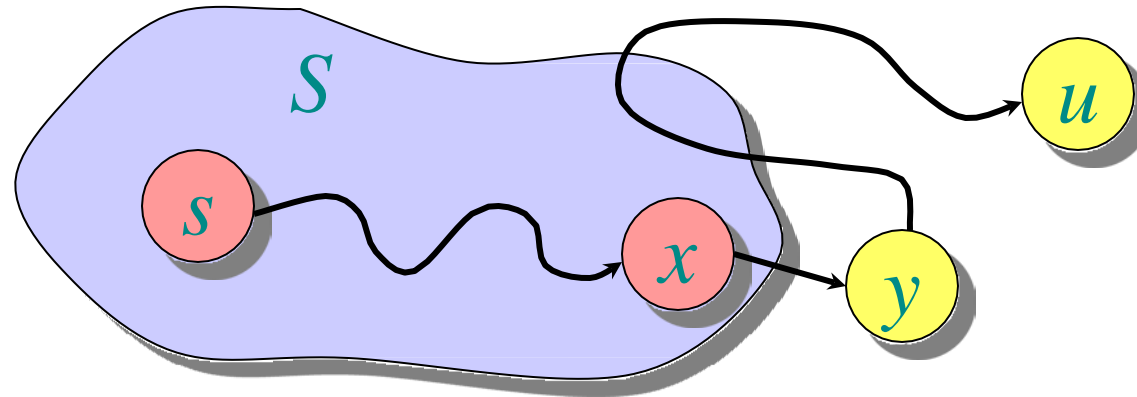
Proof.

Observe that $\delta(s, v) = \delta(s, u) + w(u, v)$. Suppose that $d[v] > \delta(s, v)$ before the relaxation. (Otherwise, we're done.) Then, the test $d[v] > d[u] + w(u, v)$ succeeds, because $d[v] > \delta(s, v) = \delta(s, u) + w(u, v) = d[u] + w(u, v)$, and the algorithm sets $d[v] = d[u] + w(u, v) = \delta(s, v)$. 

Theorem. Dijkstra's algorithm terminates with $d[v] = \delta(s, v)$ for all $v \in V$.

Proof. It suffices to show that $d[v] = \delta(s, v)$ for every $v \in V$ when v is added to S . Suppose u is the first vertex added to S for which $d[u] > \delta(s, u)$. Let y be the first vertex in $V - S$ along a shortest path from s to u , and let x be its predecessor:





Since u is the first vertex violating the claimed invariant, we have $d[x] = \delta(s, x)$.

When x was added to S , the edge (x, y) was relaxed, which implies that $d[y] = \delta(s, y) \leq \delta(s, u) < d[u]$. But, $d[u] \leq d[y]$ by our choice of u .

Contradiction.



$|V|$
times {

```
while  $Q \neq \emptyset$ 
do  $u \leftarrow \text{EXTRACT-MIN}(Q)$ 
   $S \leftarrow S \cup \{u\}$ 
  for each  $v \in \text{Adj}[u]$ 
  do if  $d[v] > d[u] + w(u, v)$ 
    then  $d[v] \leftarrow d[u] + w(u, v)$ 
```

$|V|$
times { $\text{degree}(u)$
times {

```
while  $Q \neq \emptyset$ 
do  $u \leftarrow \text{EXTRACT-MIN}(Q)$ 
   $S \leftarrow S \cup \{u\}$ 
  for each  $v \in \text{Adj}[u]$ 
    do if  $d[v] > d[u] + w(u, v)$ 
      then  $d[v] \leftarrow d[u] + w(u, v)$ 
```

$|V|$ times { **while** $Q \neq \emptyset$
do $u \leftarrow \text{EXTRACT-MIN}(Q)$
 $S \leftarrow S \cup \{u\}$
 $\text{degree}(u)$ times { **for each** $v \in \text{Adj}[u]$
do **if** $d[v] > d[u] + w(u, v)$
then $d[v] \leftarrow d[u] + w(u, v)$

Handshaking Lemma $\Rightarrow \Theta(E)$ implicit DECREASE-KEY's.

$|V|$ times { **while** $Q \neq \emptyset$
do $u \leftarrow \text{EXTRACT-MIN}(Q)$
 $S \leftarrow S \cup \{u\}$
 $\text{degree}(u)$ times { **for each** $v \in \text{Adj}[u]$
do **if** $d[v] > d[u] + w(u, v)$
then $d[v] \leftarrow d[u] + w(u, v)$

Handshaking Lemma $\Rightarrow \Theta(|E|)$ implicit DECREASE-KEY's.

$$\Theta(|V| \cdot T_{\text{EXTRACT-MIN}} + |E| \cdot T_{\text{DECREASE-KEY}})$$

Note: Same formula as in the analysis of Prim's minimum spanning tree algorithm.

$$\text{Time} = \Theta(|V|) \cdot T_{\text{EXTRACT-MIN}} + \Theta(|E|) \cdot T_{\text{DECREASE-KEY}}$$

Q	$T_{\text{EXTRACT-MIN}}$	$T_{\text{DECREASE-KEY}}$	Total
array	$O(V)$	$O(1)$	$O(V ^2)$

$$\text{Time} = \Theta(|V|) \cdot T_{\text{EXTRACT-MIN}} + \Theta(|E|) \cdot T_{\text{DECREASE-KEY}}$$

Q	$T_{\text{EXTRACT-MIN}}$	$T_{\text{DECREASE-KEY}}$	Total
array	$O(V)$	$O(1)$	$O(V ^2)$
binary heap	$O(\lg V)$	$O(\lg V)$	$O(E \lg V)$

$$\text{Time} = \Theta(|V|) \cdot T_{\text{EXTRACT-MIN}} + \Theta(|E|) \cdot T_{\text{DECREASE-KEY}}$$

Q	$T_{\text{EXTRACT-MIN}}$	$T_{\text{DECREASE-KEY}}$	Total
array	$O(V)$	$O(1)$	$O(V ^2)$
binary heap	$O(\lg V)$	$O(\lg V)$	$O(E \lg V)$
Fibonacci heap	$O(\lg V)$ amortized	$O(1)$ amortized	$O(E + V \lg V)$ worst case

Suppose that $w(u, v) = 1$ for all $(u, v) \in E$. Can Dijkstra's algorithm be improved?

- Use a simple FIFO queue instead of a priority queue.

Suppose that $w(u, v) = 1$ for all $(u, v) \in E$. Can Dijkstra's algorithm be improved?

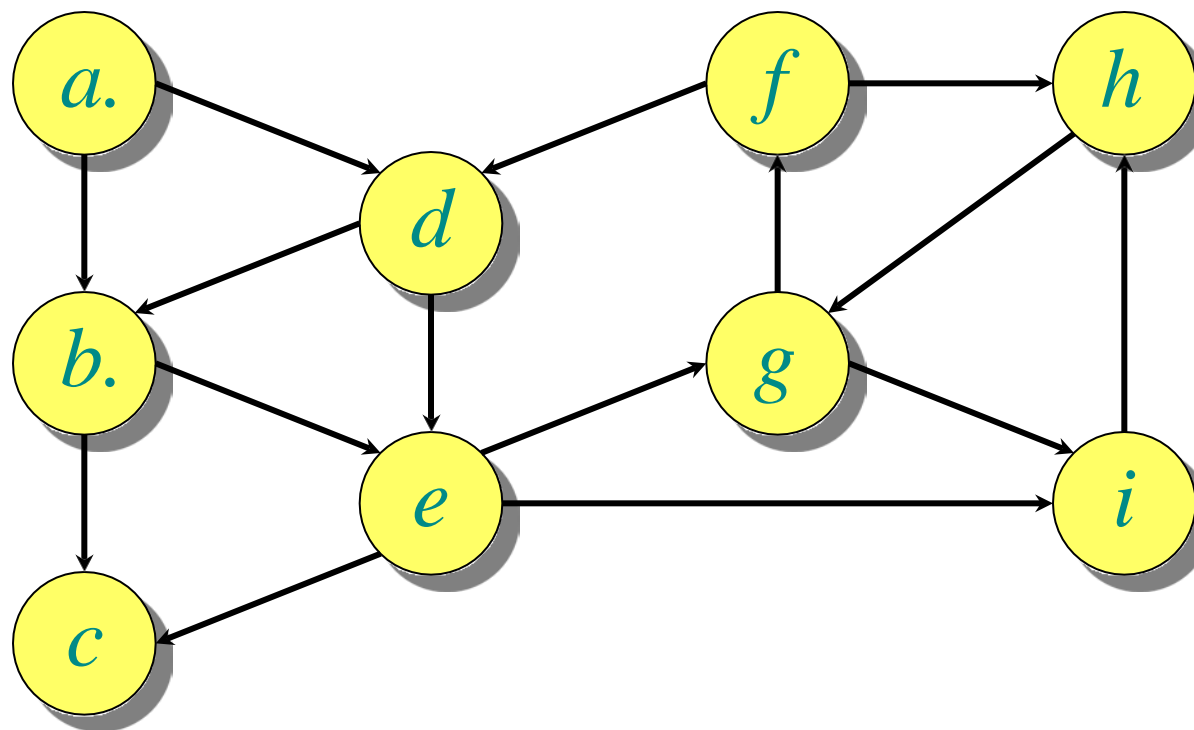
- Use a simple FIFO queue instead of a priority queue.

Breadth-first search

```
while  $Q \neq \emptyset$ 
  do  $u \leftarrow \text{DEQUEUE}(Q)$ 
    for each  $v \in \text{Adj}[u]$ 
      do if  $d[v] = \infty$ 
        then  $d[v] \leftarrow d[u] + 1$ 
           ENQUEUE( $Q, v$ )
```

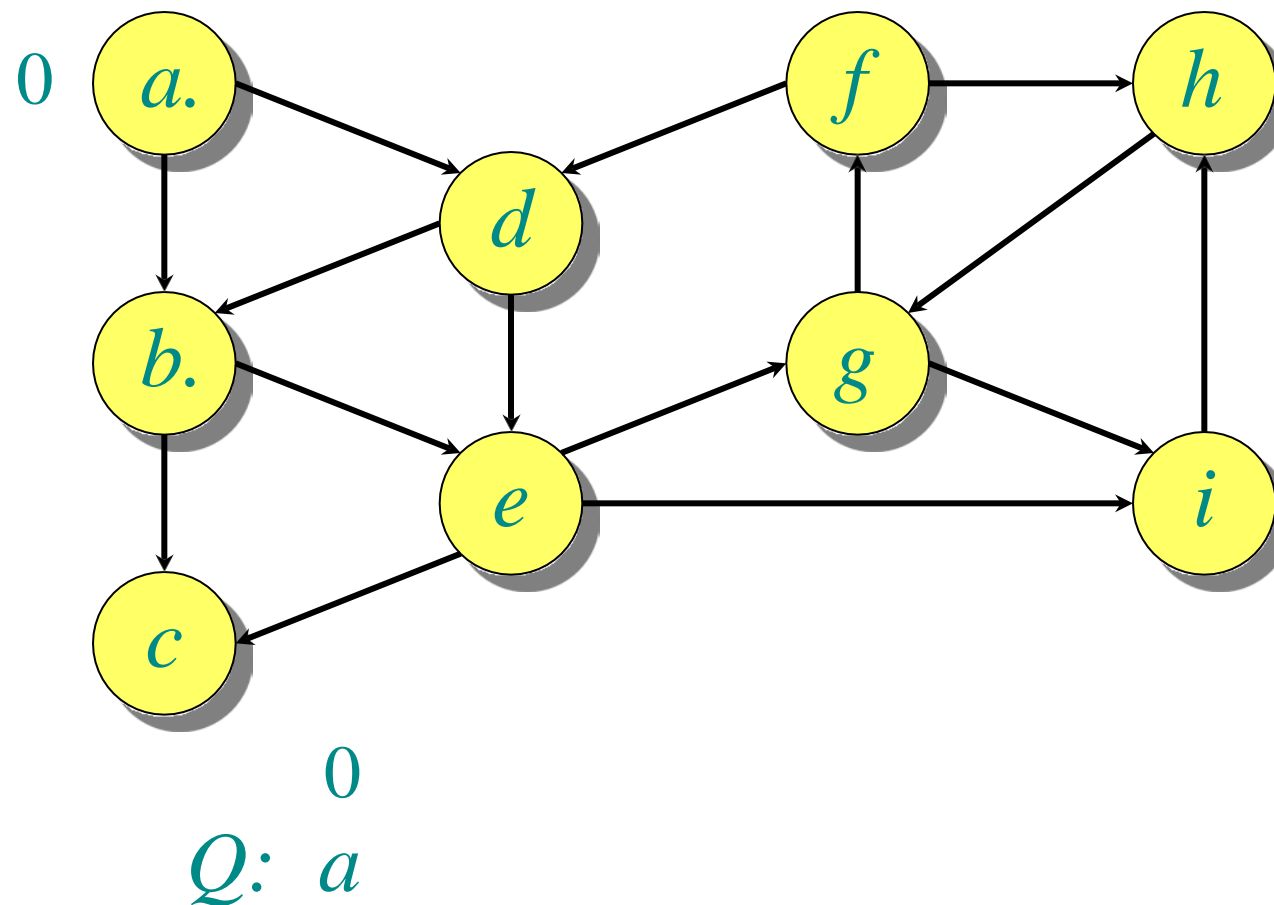
Analysis: Time = $O(|V| + |E|)$.

Example of breadth-first search

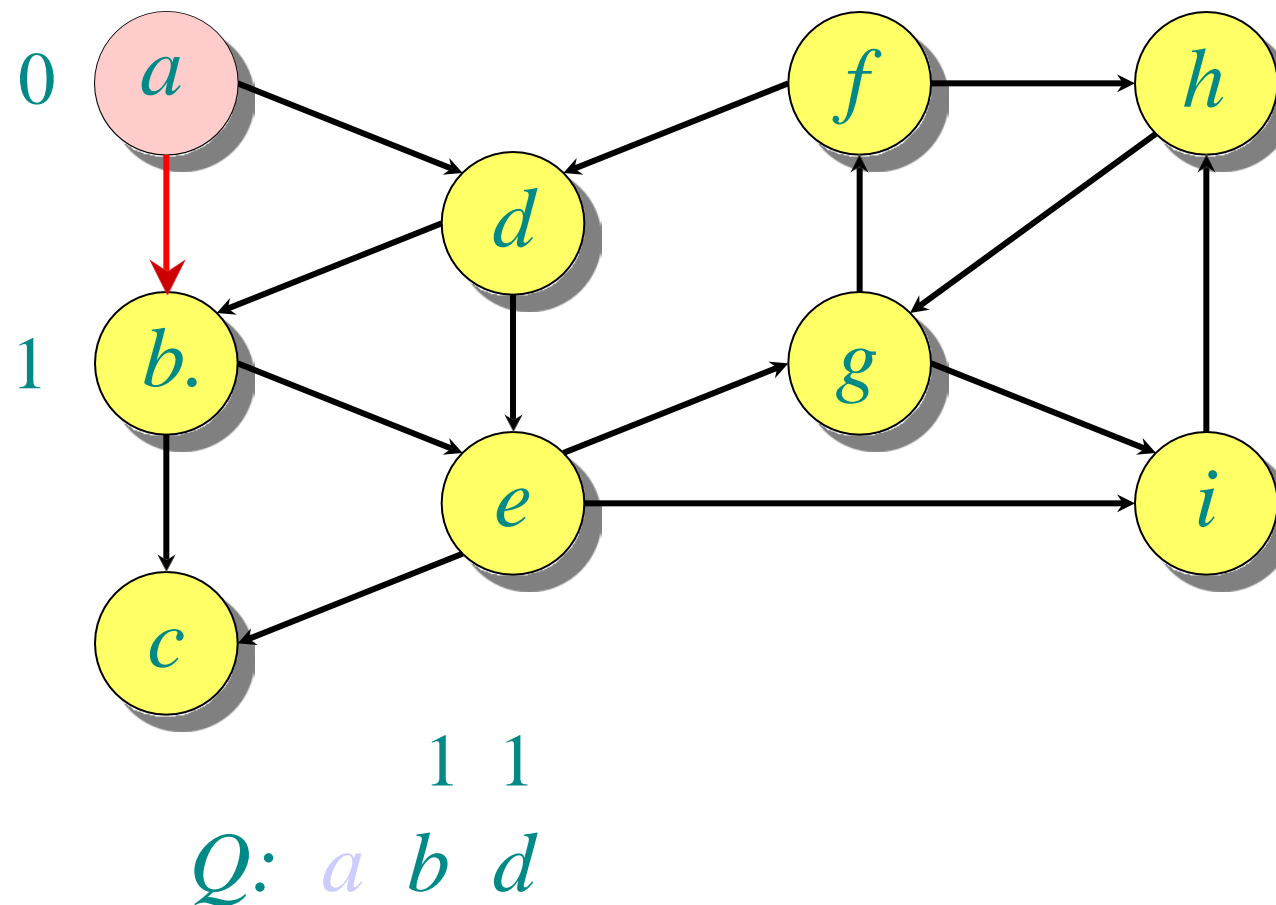


Q :

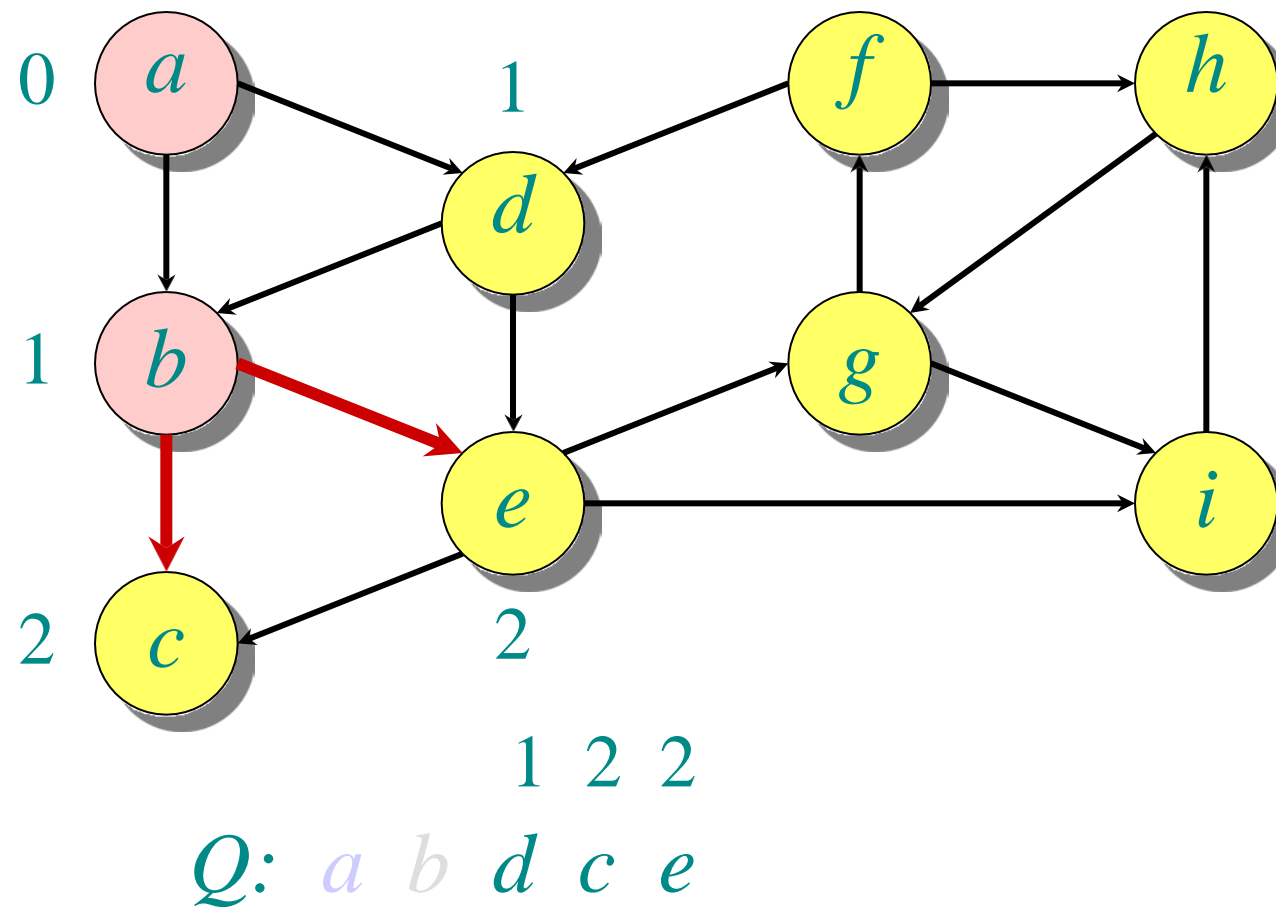
Example of breadth-first search



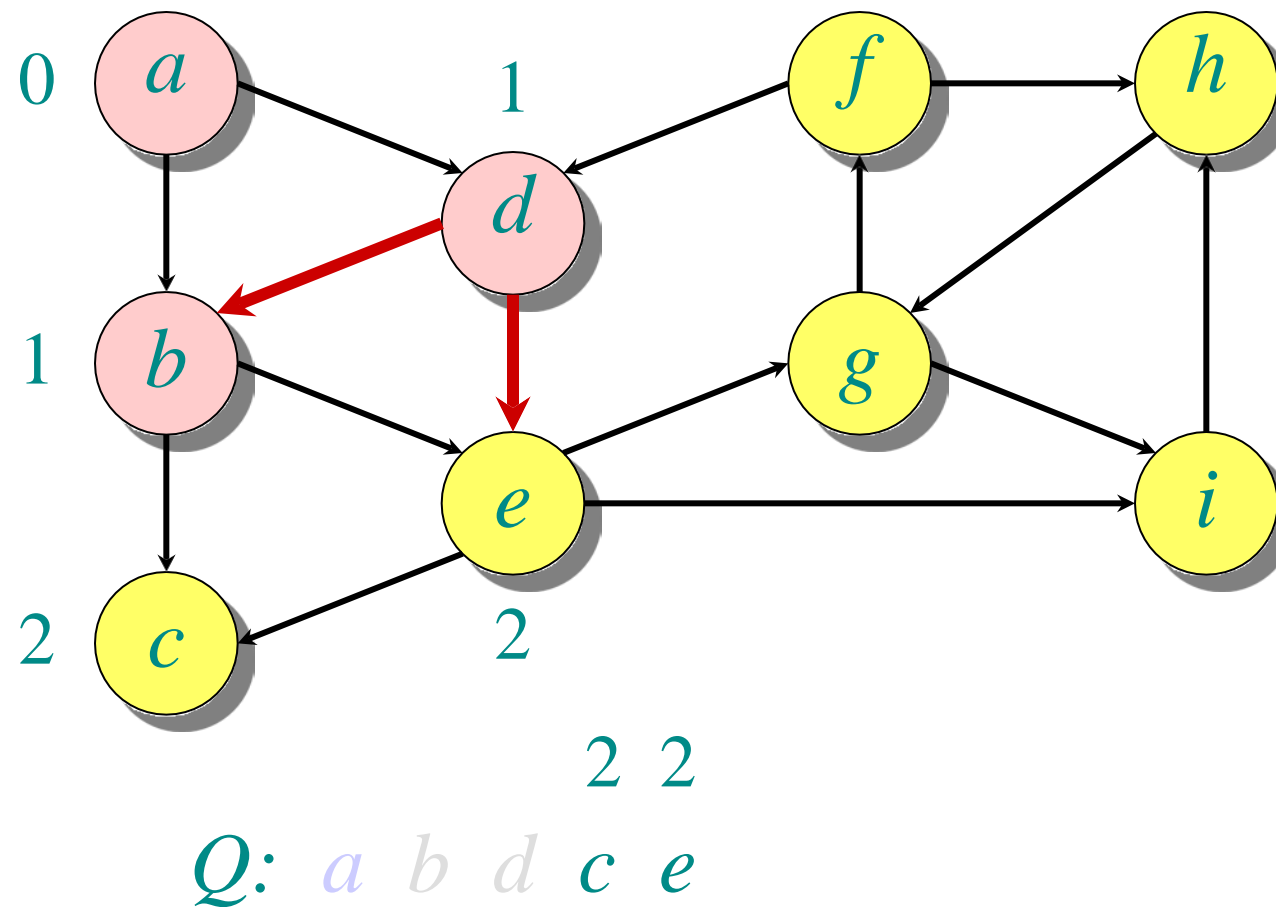
Example of breadth-first search



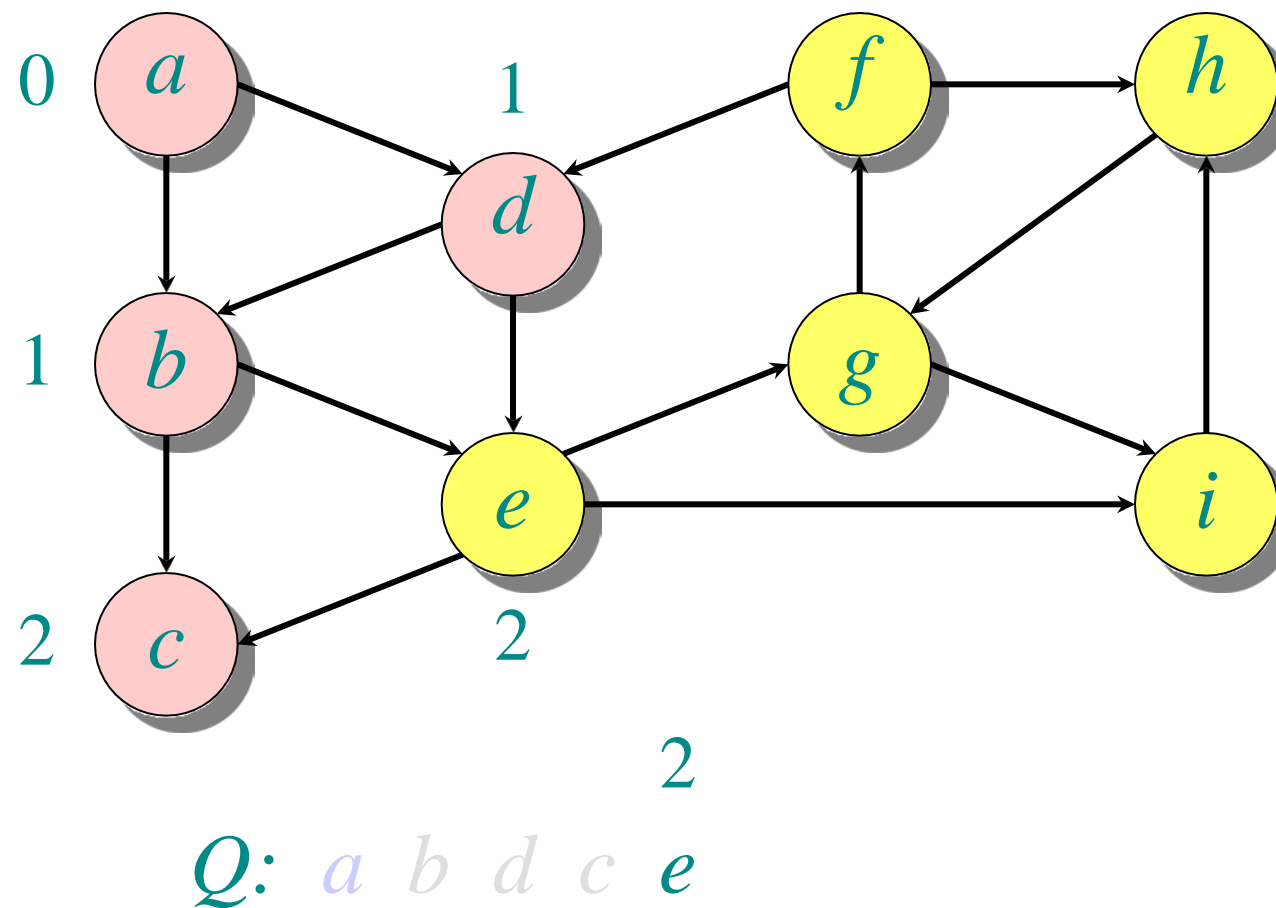
Example of breadth-first search



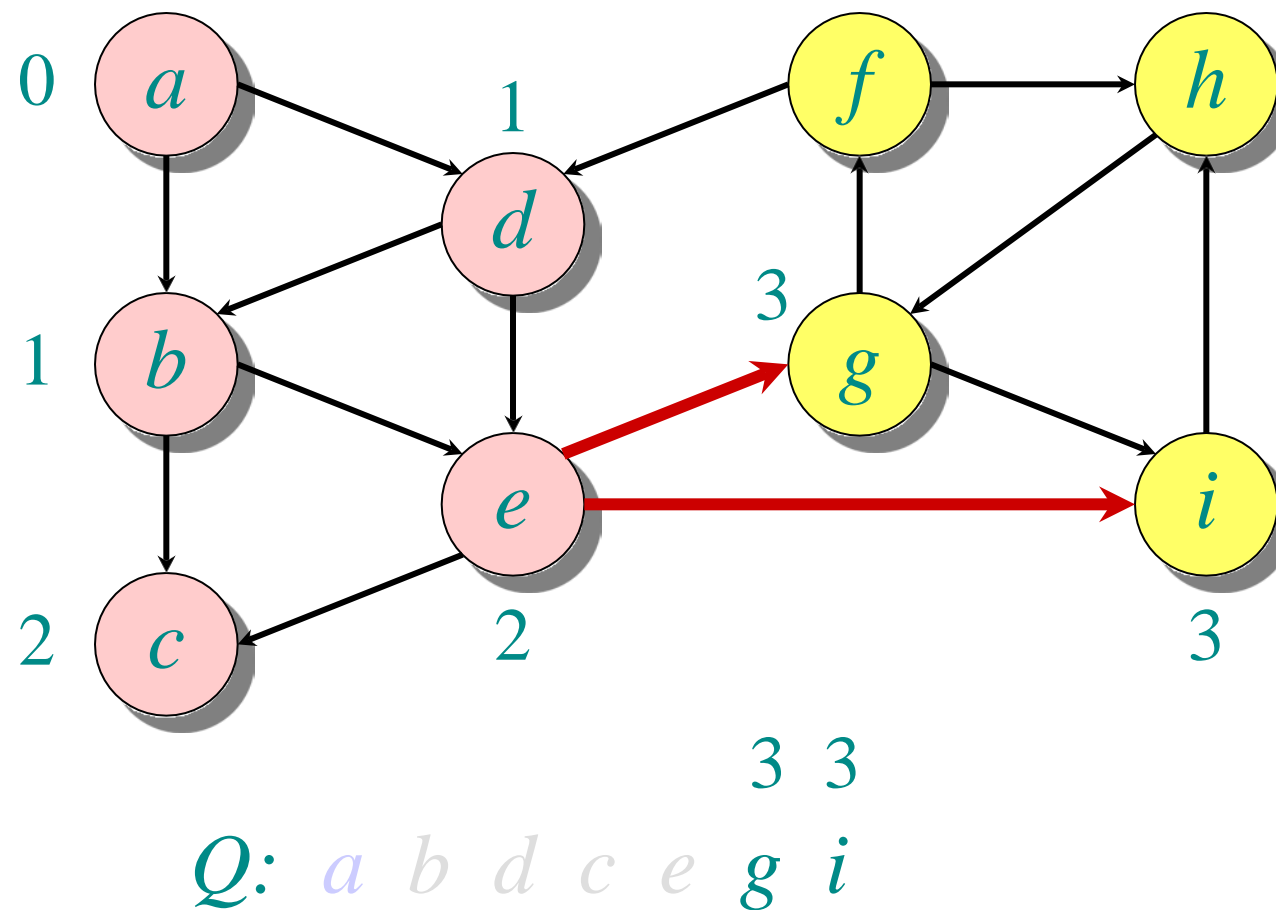
Example of breadth-first search



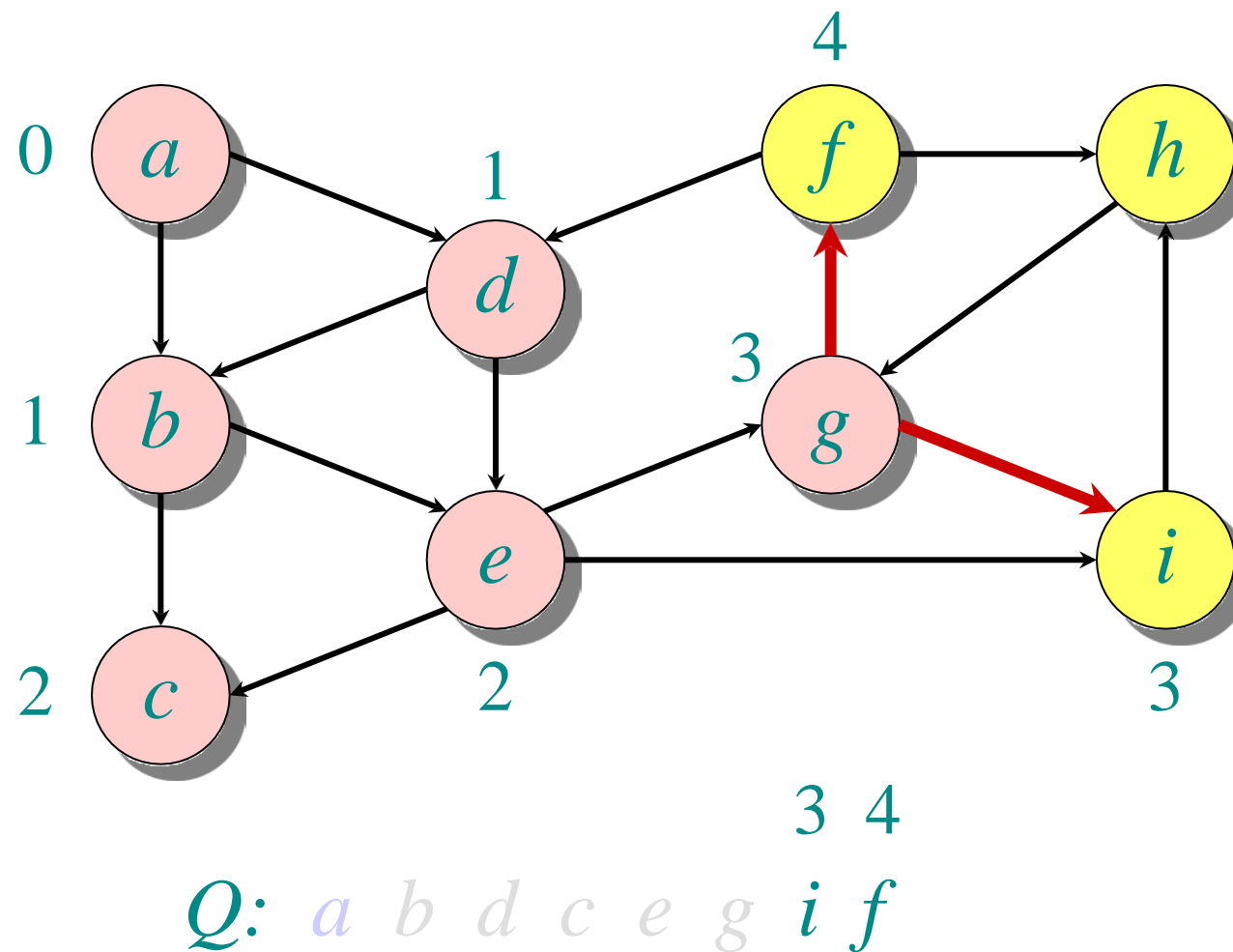
Example of breadth-first search



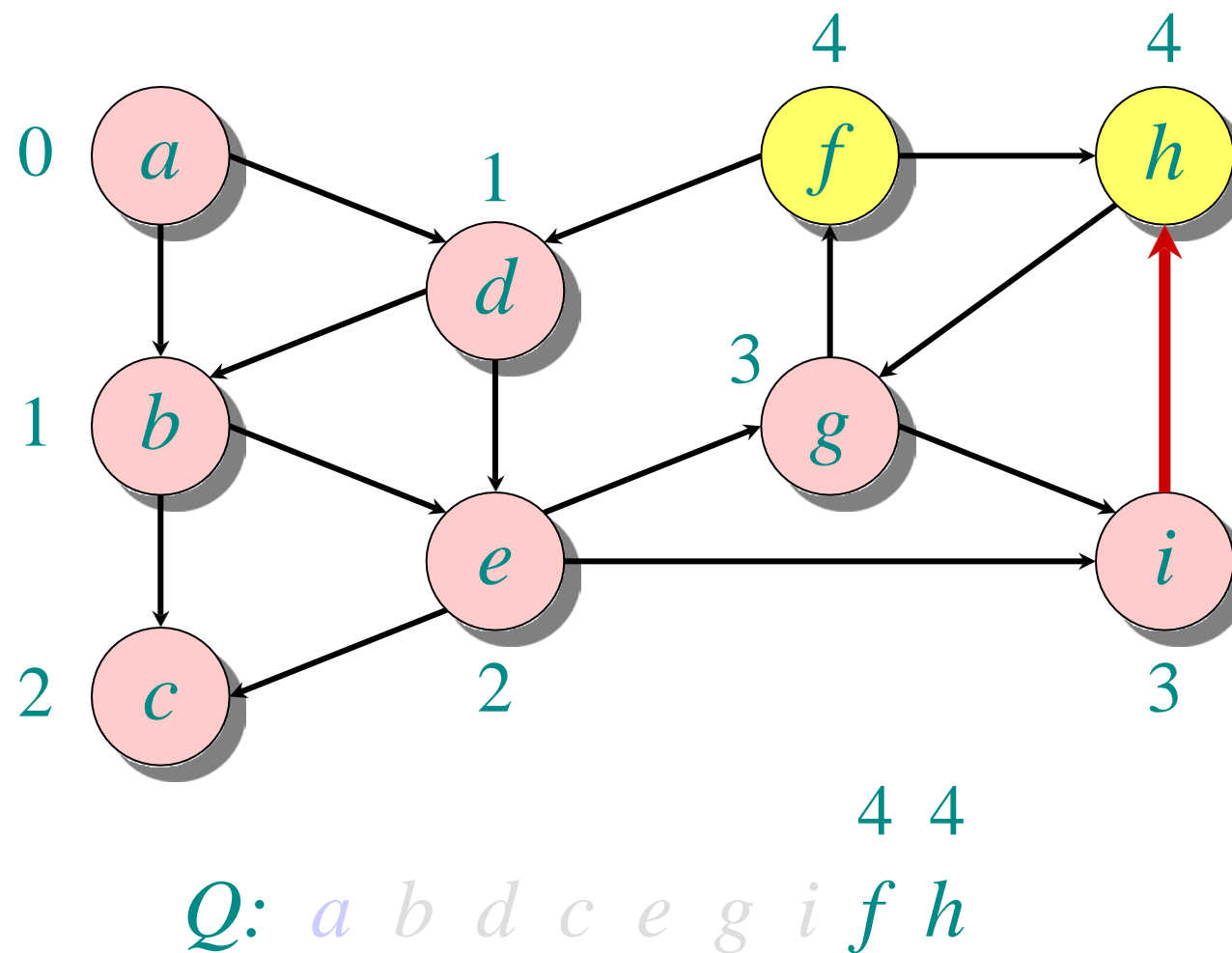
Example of breadth-first search



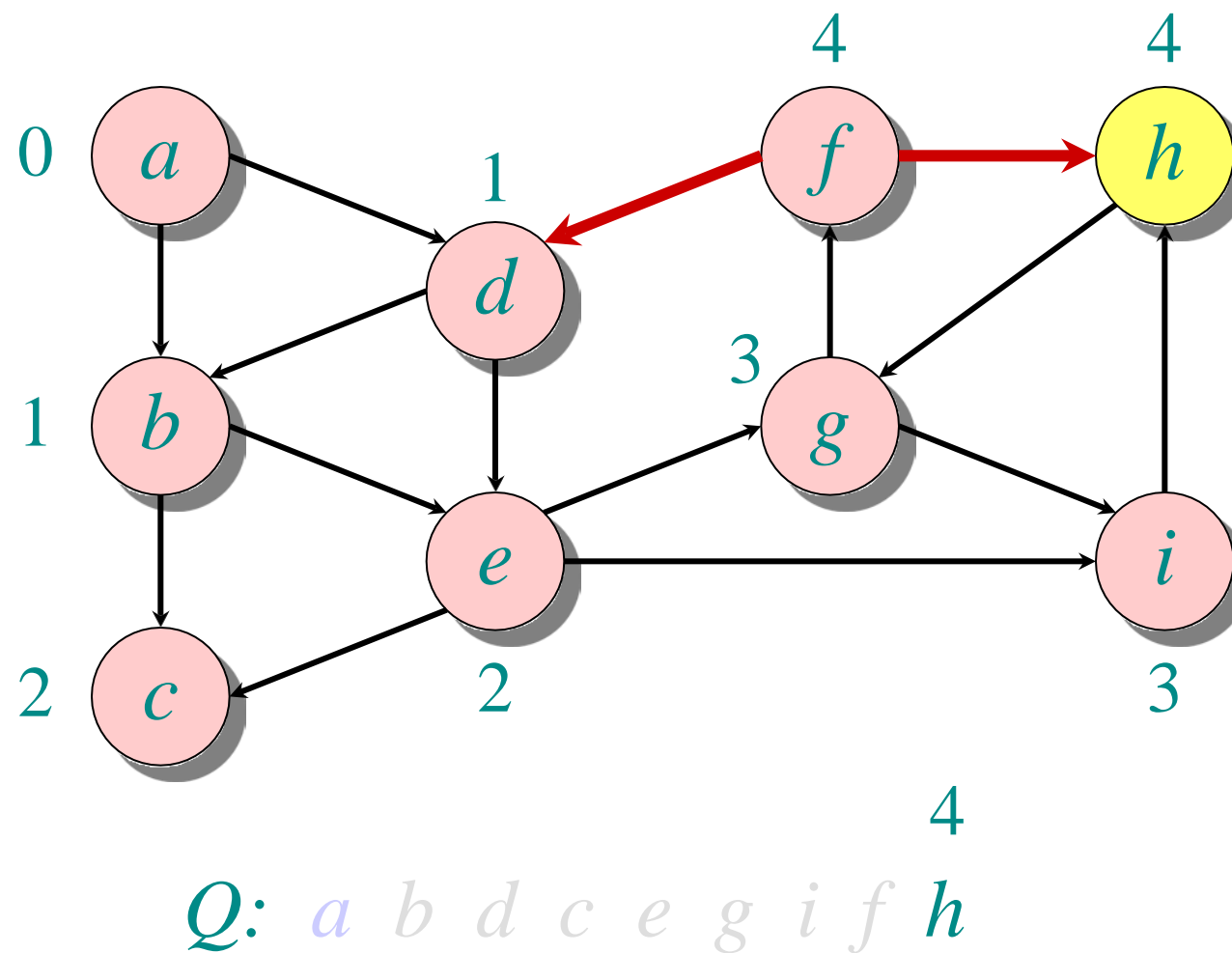
Example of breadth-first search



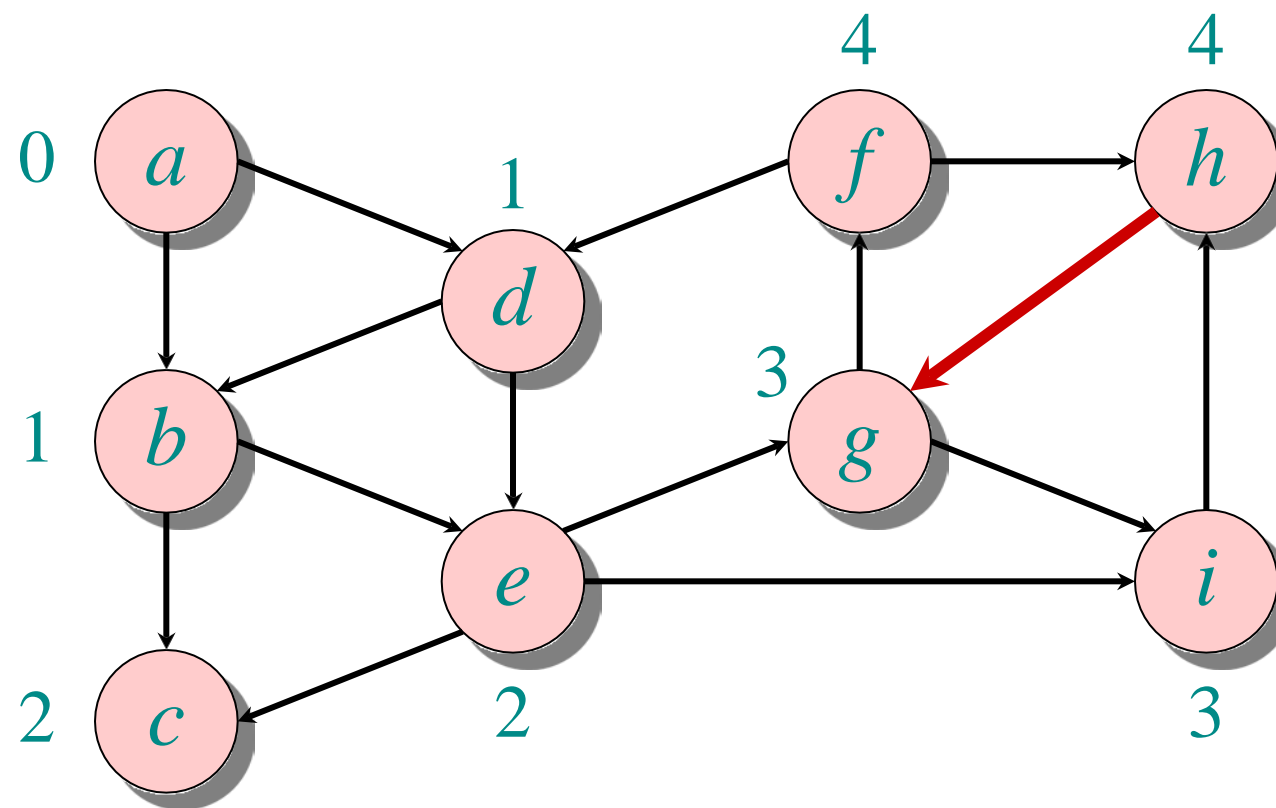
Example of breadth-first search



Example of breadth-first search

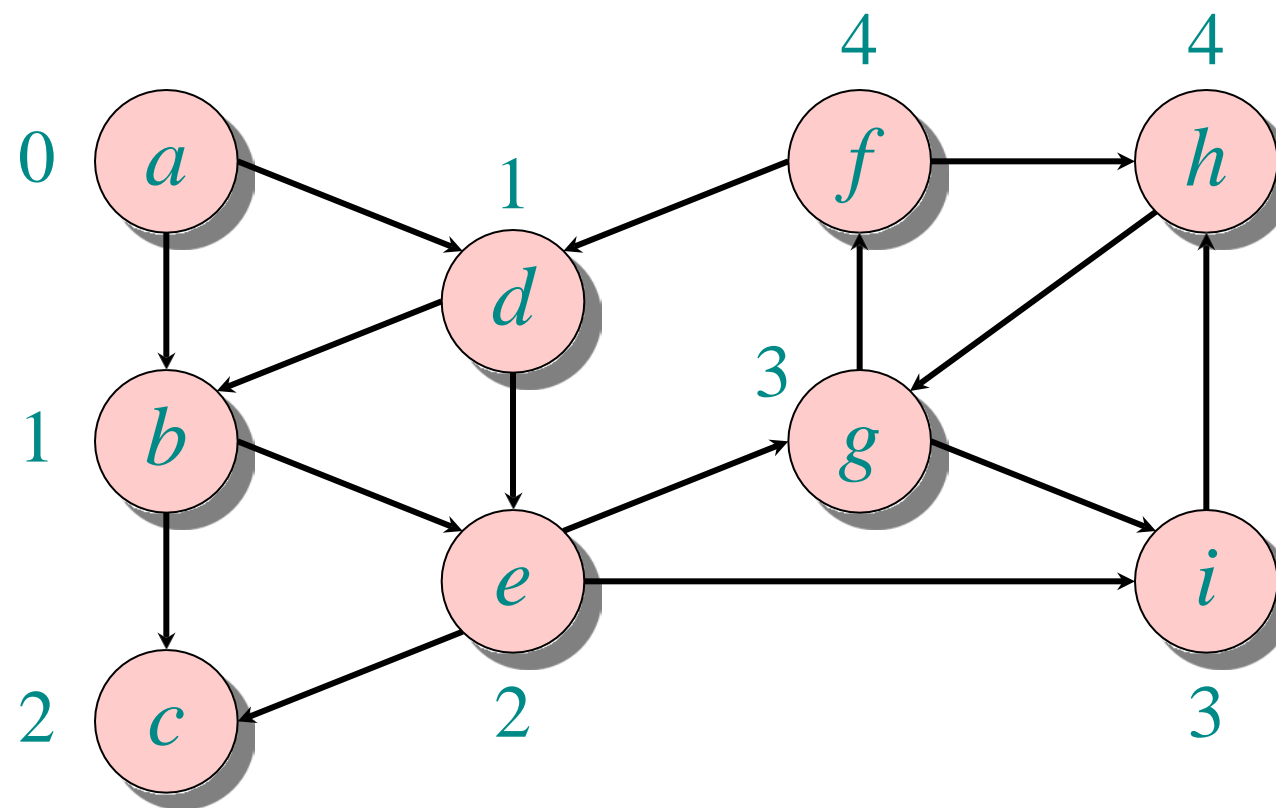


Example of breadth-first search



Q: *a* *b* *d* *c* *e* *g* *i* *f* *h*

Example of breadth-first search



Q: *a b d c e g i f h*


```
while  $Q \neq \emptyset$ 
  do  $u \leftarrow \text{DEQUEUE}(Q)$ 
    for each  $v \in \text{Adj}[u]$ 
      do if  $d[v] = \infty$ 
        then  $d[v] \leftarrow d[u] + 1$ 
          ENQUEUE( $Q, v$ )
```

Key idea:

The FIFO Q in breadth-first search mimics the priority queue Q in Dijkstra.

- **Invariant:** v comes after u in Q implies that $d[v] = d[u]$ or $d[v] = d[u] + 1$.

Bellman-Ford algorithm

- (-) Slower than Dijkstra's algorithm
- (+) Can handle negative edge weights.
 - Can be useful if you want to say that some edges are actively good to take, rather than costly.
 - Can be useful as a building block in other algorithms.

Basic idea:

Instead of picking the u with the smallest $d[u]$ to update, just update all of the u 's simultaneously.

Bellman-Ford(G, s):

- $d[v] = \infty$ for all v in V
 - $d[s] = 0$
 - **For** $i=0, \dots, |V|-1$:
 - **For** u in V :
 - **For** v in $u.\text{neighbors}$:
 - $d[v] \leftarrow \min(d[v], d[u] + \text{edgeWeight}(u, v))$
- Instead of picking u cleverly,
just update for all of the u 's.

Compare to Dijkstra:

- **While** there are **not-sure** nodes:
 - Pick the **not-sure** node u with the smallest estimate $d[u]$.
 - **For** v in $u.\text{neighbors}$:
 - $d[v] \leftarrow \min(d[v], d[u] + \text{edgeWeight}(u, v))$
 - Mark u as **sure**.

- We are actually going to change this to be less smart.
- Keep n arrays: $d^{(0)}, d^{(1)}, \dots, d^{(n-1)}$

Bellman-Ford*(G,s):

- $d^{(i)}[v] = \infty$ for all v in V , for all $i=0, \dots, |V|-1$
- $d^{(0)}[s] = 0$
- **For** $i=0, \dots, |V|-2$:
 - **For** u in V :
 - **For** v in u .neighbors:
 - $d^{(i+1)}[v] \leftarrow \min(d^{(i)}[v], d^{(i+1)}[v], d^{(i)}[u] + \text{edgeWeight}(u,v))$
- Then $\text{dist}(s,v) = d^{(n-1)}[v]$

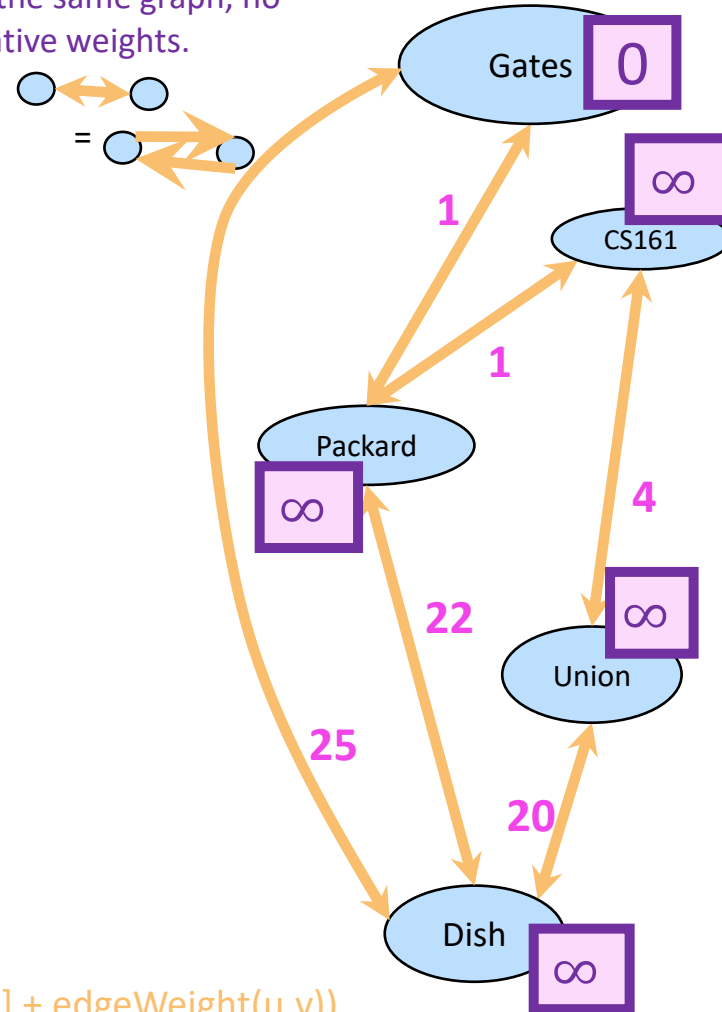
Slightly different than the original Bellman-Ford algorithm, but the analysis is basically the same.

How far is a node from Gates?

	Gates	Packard	CS161	Union	Dish
$d^{(0)}$	0	∞	∞	∞	∞
$d^{(1)}$					
$d^{(2)}$					
$d^{(3)}$					
$d^{(4)}$					

- For $i=0, \dots, |V|-2$:
 - For u in V :
 - For v in $u.\text{neighbors}$:
 - $d^{(i+1)}[v] \leftarrow \min(d^{(i)}[v], d^{(i+1)}[v], d^{(i)}[u] + \text{edgeWeight}(u,v))$

Start with the same graph, no negative weights.

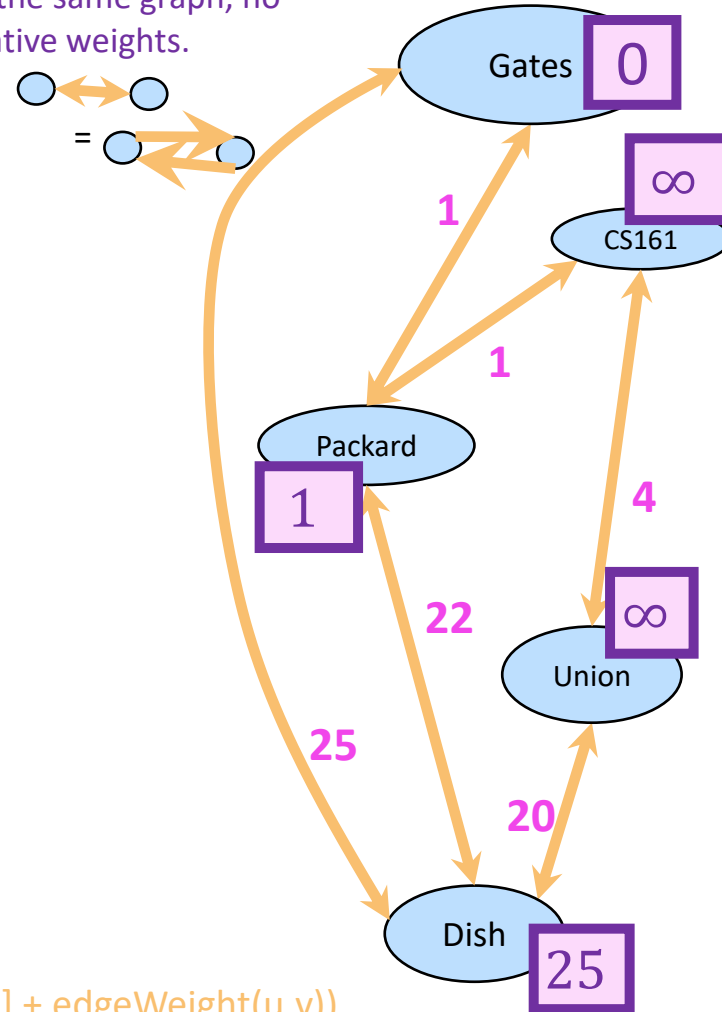


How far is a node from Gates?

	Gates	Packard	CS161	Union	Dish
$d^{(0)}$	0	∞	∞	∞	∞
$d^{(1)}$	0	1	∞	∞	25
$d^{(2)}$					
$d^{(3)}$					
$d^{(4)}$					

- For $i=0, \dots, |V|-2$:
 - For u in V :
 - For v in $u.\text{neighbors}$:
 - $d^{(i+1)}[v] \leftarrow \min(d^{(i)}[v], d^{(i+1)}[v], d^{(i)}[u] + \text{edgeWeight}(u,v))$

Start with the same graph, no negative weights.

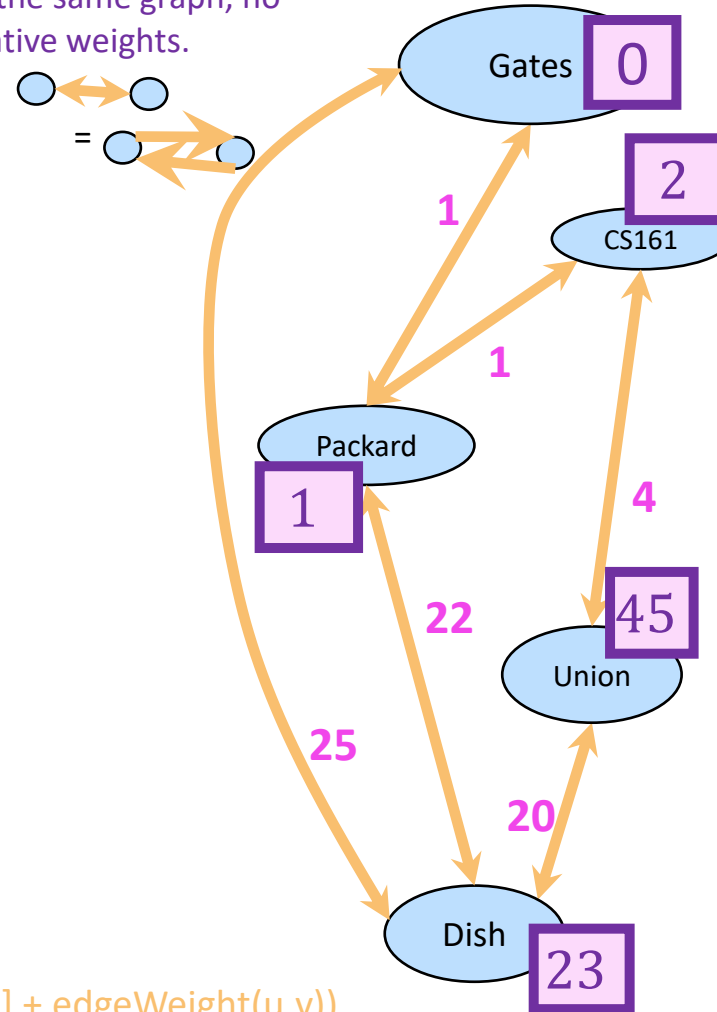


How far is a node from Gates?

	Gates	Packard	CS161	Union	Dish
$d^{(0)}$	0	∞	∞	∞	∞
$d^{(1)}$	0	1	∞	∞	25
$d^{(2)}$	0	1	2	45	23
$d^{(3)}$					
$d^{(4)}$					

- For $i=0, \dots, |V|-2$:
 - For u in V :
 - For v in $u.\text{neighbors}$:
 - $d^{(i+1)}[v] \leftarrow \min(d^{(i)}[v], d^{(i+1)}[v], d^{(i)}[u] + \text{edgeWeight}(u,v))$

Start with the same graph, no negative weights.

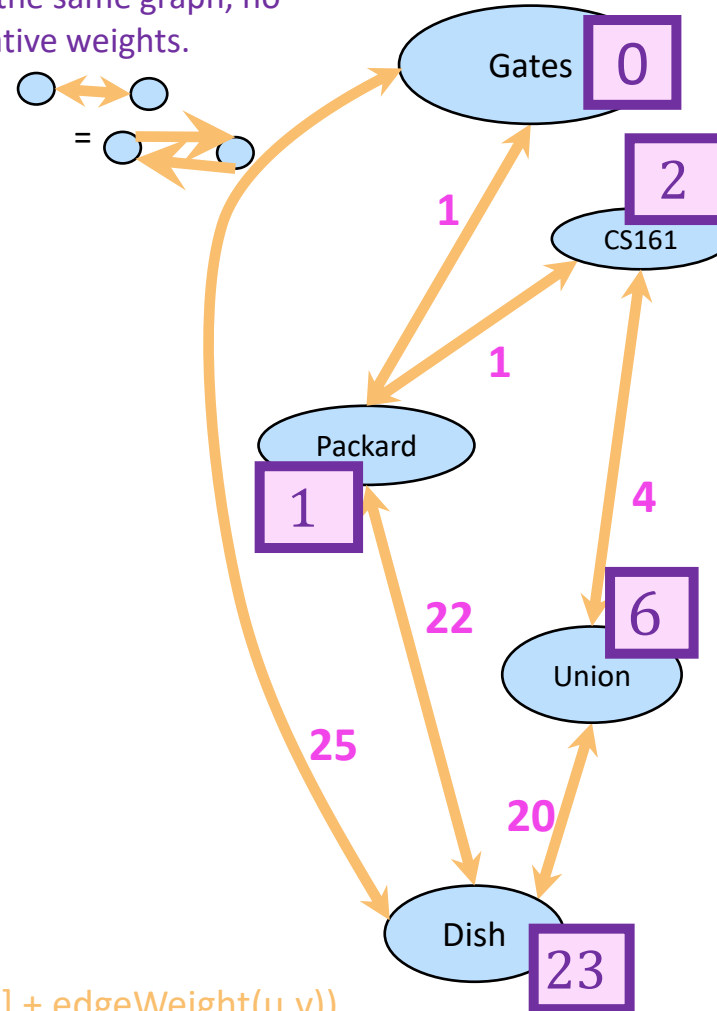


How far is a node from Gates?

	Gates	Packard	CS161	Union	Dish
$d^{(0)}$	0	∞	∞	∞	∞
$d^{(1)}$	0	1	∞	∞	25
$d^{(2)}$	0	1	2	45	23
$d^{(3)}$	0	1	2	6	23
$d^{(4)}$					

- For $i=0, \dots, |V|-2$:
 - For u in V :
 - For v in $u.\text{neighbors}$:
 - $d^{(i+1)}[v] \leftarrow \min(d^{(i)}[v], d^{(i+1)}[v], d^{(i)}[u] + \text{edgeWeight}(u,v))$

Start with the same graph, no negative weights.



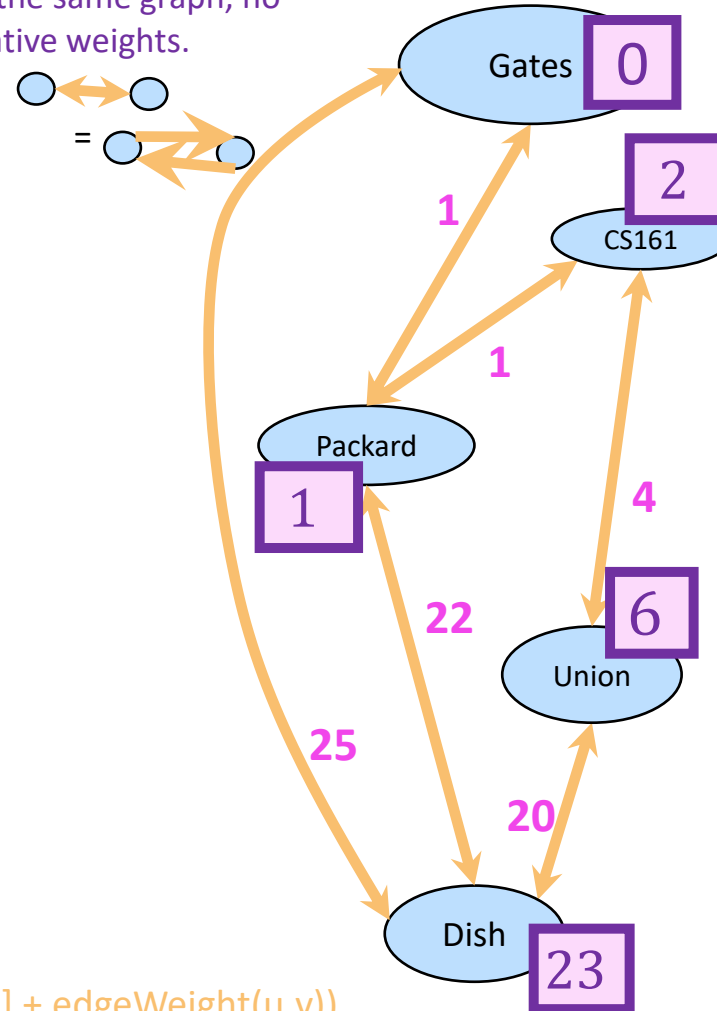
How far is a node from Gates?

	Gates	Packard	CS161	Union	Dish
$d^{(0)}$	0	∞	∞	∞	∞
$d^{(1)}$	0	1	∞	∞	25
$d^{(2)}$	0	1	2	45	23
$d^{(3)}$	0	1	2	6	23
$d^{(4)}$	0	1	2	6	23

These are the final distances!

- For $i=0, \dots, |V|-2$:
 - For u in V :
 - For v in $u.\text{neighbors}$:
 - $d^{(i+1)}[v] \leftarrow \min(d^{(i)}[v], d^{(i+1)}[v], d^{(i)}[u] + \text{edgeWeight}(u,v))$

Start with the same graph, no negative weights.



- Does it work?

- Yes

- Idea to the right.

- Is it fast?

- Not really...



A **simple path** is a path with no cycles.

	Gates	Packard	CS161	Union	Dish
$d^{(0)}$	0	∞	∞	∞	∞
$d^{(1)}$	0	1	∞	∞	25
$d^{(2)}$	0	1	2	45	23
$d^{(3)}$	0	1	2	6	23
$d^{(4)}$	0	1	2	6	23

Inductive Hypothesis:

$d^{(i)}[v]$ is equal to the cost of the shortest path between s and v **with at most i edges**.

Conclusion:

$d^{(|V|-1)}[v]$ is equal to the cost of the shortest simple path between s and v . **(Since all simple paths have at most $|V|-1$ edges).**

- **Inductive Hypothesis:**

- After iteration i , for each v , $d^{(i)}[v]$ is equal to the cost of the shortest path between s and v with at most i edges.

- **Base case:**

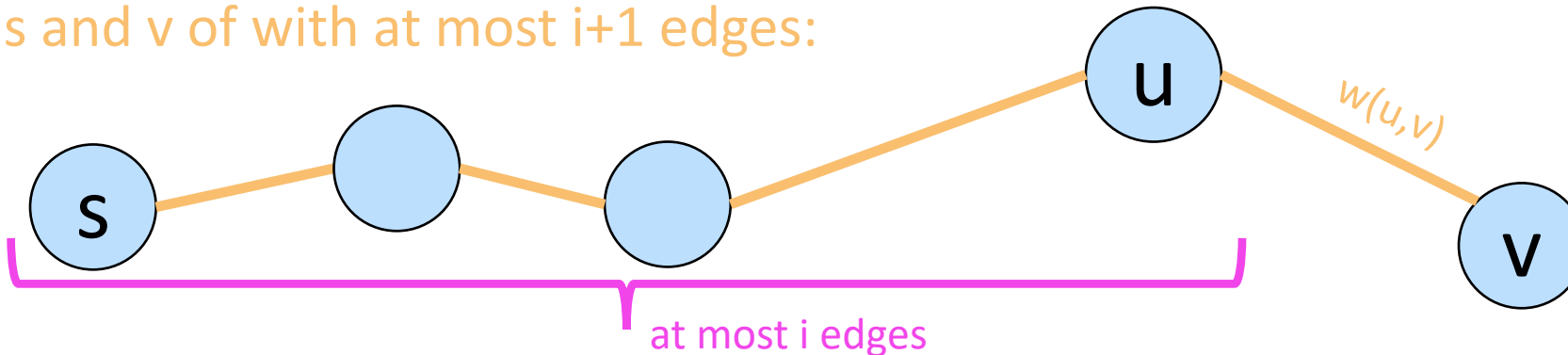
- After iteration 0...



- **Inductive step:**

- Suppose the inductive hypothesis holds for i .
- We want to establish it for $i+1$.

Say this is the shortest path between s and v of with at most $i+1$ edges:



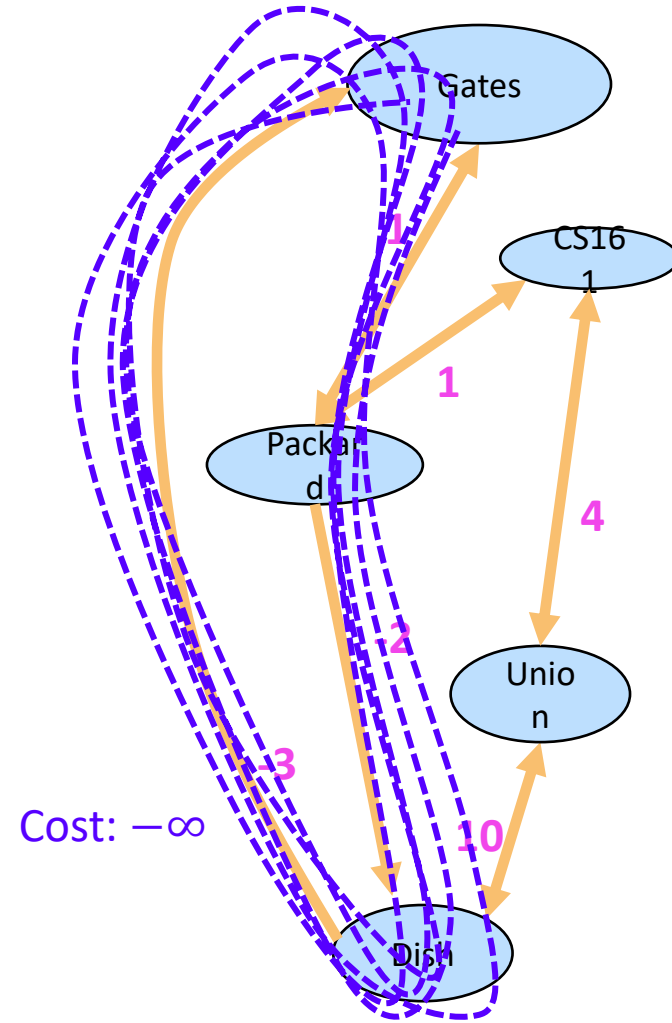
Hypothesis: After iteration i , for each v , $d^{(i)}[v]$ is equal to the cost of the shortest path between s and v with at most i edges.

- By induction, $d^{(i)}[u]$ is the cost of a shortest path between s and u of i edges.
- By setup, $d^{(i)}[u] + w(u,v)$ is the cost of a shortest path between s and v of $i+1$ edges.
- In the $i+1$ 'st iteration, we ensure $d^{(i+1)}[v] \leq d^{(i)}[u] + w(u,v)$.
- So $d^{(i+1)}[v] \leq$ cost of shortest path between s and v with $i+1$ edges.
- But $d^{(i+1)}[v] =$ cost of a particular path of at most $i+1$ edges \geq cost of shortest path.
- So $d[v] =$ cost of shortest path with at most $i+1$ edges.

- Running time: $O(|V||E|)$ running time
 - For each of $|V|$ steps we update m edges
 - Slower than Dijkstra
- However, it's also more flexible in a few ways.
 - Can handle negative edges
 - If we constantly do these iterations, any changes in the network will eventually propagate through.

Negative edge weights?

- What is the shortest path from Gates to the Union?
- Shortest paths aren't defined if there are negative cycles!
- B-F works with negative edge weights...as long as there are not negative cycles.
 - A negative cycle is a path with the same start and end vertex whose cost is negative.
- However, B-F can detect negative cycles.



How Bellman-Ford deals with negative cycles

- If there are no negative cycles:
 - Everything works as it should.
 - The algorithm stabilizes after $|V|-1$ rounds.
 - Note: Negative *edges* are okay!!
- If there are negative cycles:
 - Not everything works as it should...
 - it couldn't possibly work, since shortest paths aren't well-defined if there are negative cycles.
 - The $d[v]$ values will keep changing.
- Solution:
 - Go one round more and see if things change.
 - If so, return NEGATIVE CYCLE ☹️

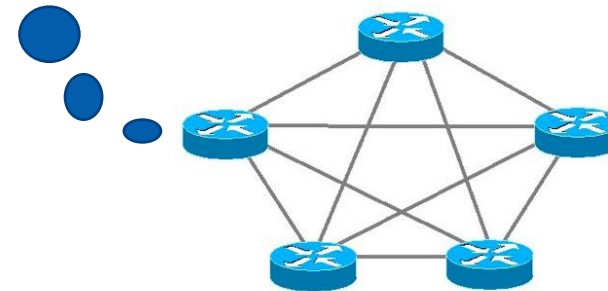
- The Bellman-Ford algorithm:
 - Finds shortest paths in weighted graphs with negative edge weights
 - runs in time $O(|V| |E|)$ on a graph G with n vertices and m edges.
- If there are no negative cycles in G :
 - the BF algorithm terminates with $d^{(|V|-1)}[v] = d(s,v)$.
- If there are negative cycles in G :
 - the BF algorithm returns negative cycle.

Bellman-Ford is also used in practice.

- eg, Routing Information Protocol (RIP) uses something like Bellman-Ford.
 - Older protocol, not used as much anymore.

- Each router keeps a **table** of distances to every other router.
- Periodically we do a Bellman-Ford update.
- This means that if there are changes in the network, this will propagate. (maybe slowly...)

Destination	Cost to get there	Send to whom?
172.16.1.0	34	172.16.1.1
10.20.40.1	10	192.168.1.2
10.155.120.1	9	10.13.50.0



Single-source shortest paths

- Nonnegative edge weights
 - ★ Dijkstra's algorithm: $O(|E| + |V| \lg |V|)$
- General
 - ★ Bellman-Ford algorithm: $O(|V||E|)$

All-pairs shortest paths

- Nonnegative edge weights
 - ★ Dijkstra's algorithm $|V|$ times: $O(|V||E| + |V|^2 \lg |V|)$
- General
 - ★ Floyd-Warshall algorithms: $\Theta(|V|^3)$.

Remain for Dynamic Programming

The End