CSE332: Data Abstractions

Lecture 14: Shortest Paths

James Fogarty

Winter 2012

Including slides developed in part by
Ruth Anderson, James Fogarty, Dan Grossman, Richard Ladner, Steve Seitz

Single Source Shortest Paths

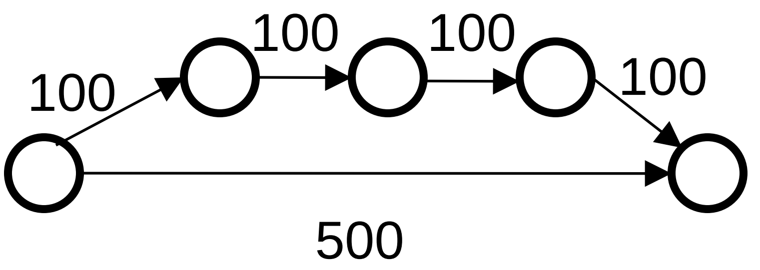
- Done: BFS for minimum path length from **v** to **u** in time $O(|E|+(|V|)$
- Actually, can find the minimum path length from **v** to *every node*
  - Still $O(|E|+(|V|)$
  - No faster way for a "distinguished" destination in the worst-case
- Now: Weighted graphs

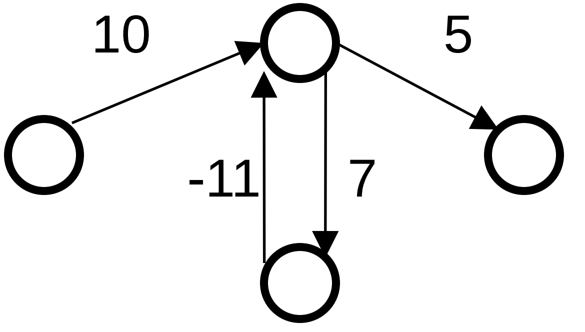
Given a weighted graph and node **v**,

find the minimum-cost path from **v** to every node

- As before, asymptotically no harder than for one destination
- Unlike before, BFS will not work

Not as Easy





Why BFS won't work: Shortest path may not have the fewest edges

– Annoying when this happens with costs of flights

We will assume there are no negative weights

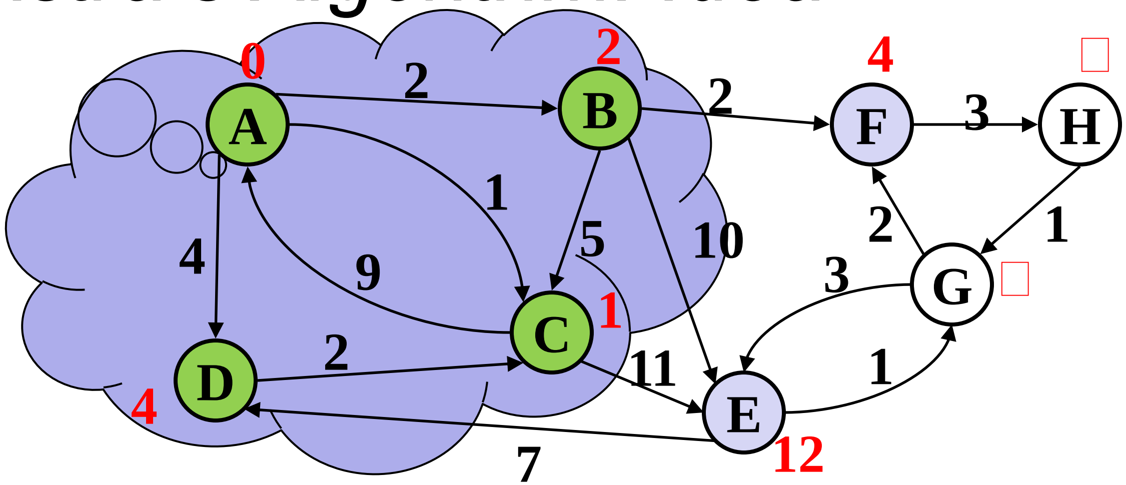
- Problem is ill-defined if there are negative-cost *cycles*
- Today's algorithm is wrong if *edges* can be negative

Dijkstra's Algorithm

- Named after its inventor Edsger Dijkstra (1930-2002)
  - Truly one of the "founders" of computer science; this is just one of his many contributions
  - Sample quotation: "computer science is no more about computers than astronomy is about telescopes"

- The idea: reminiscent of BFS, but adapted to handle weights
  - A priority queue will prove useful for efficiency
  - Grow set of nodes whose shortest distance has been computed
  - Nodes not in the set will have a "best distance so far"

Dijkstra's Algorithm: Idea



- Initially, start node has cost 0 and all other nodes have cost ∞
- At each step:
  - Pick closest unknown vertex **v**
  - Add it to the "cloud" of known vertices
  - Update distances for nodes with edges from **v**
- That's it!   But we need to prove it produces correct answers

The Algorithm

1. For each node **`v`**, set **`v.cost =`** ∞ and **`v.known = false`**
2. Set **`source.cost = 0`**
3. While there are unknown nodes in the graph
   a) Select the unknown node **`v`** with lowest cost
   b) Mark **`v`** as known
   c) For each edge **`(v,u)`** with weight **`w`**,

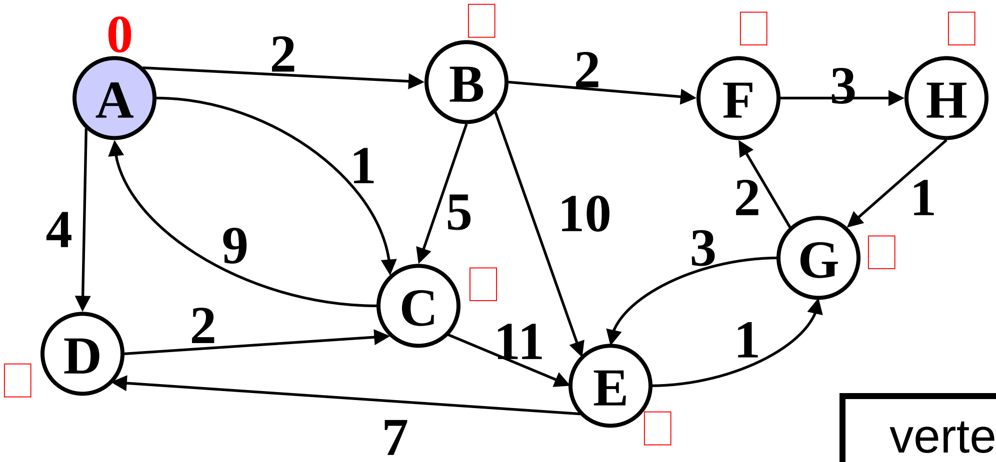
```
c1 = v.cost + w // cost of best path through v to u
c2 = u.cost   // cost of best path to u previously known
if(c1 < c2){ // if the path through v is better
  u.cost = c1
  u.path = v // for computing actual paths
}
```

Important Features

- When a vertex is marked known,
  the cost of the shortest path to that node is known
  - The path is also known by following back-pointers

- While a vertex is still not known,
  another shorter path to it might still be found

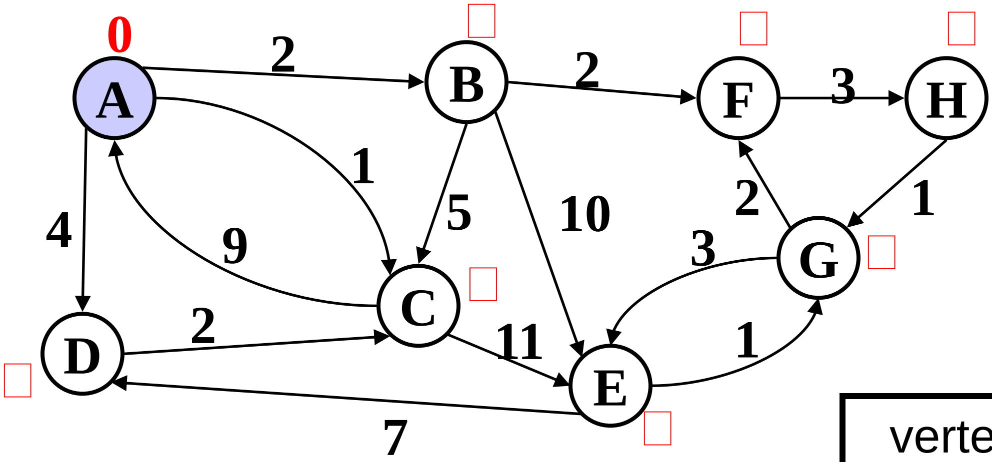
Example #1



vertex	known?	cost	path
A			
B			
C			
D			
E			
F			
G			
H			

<u>Order Added to Known Set:</u>

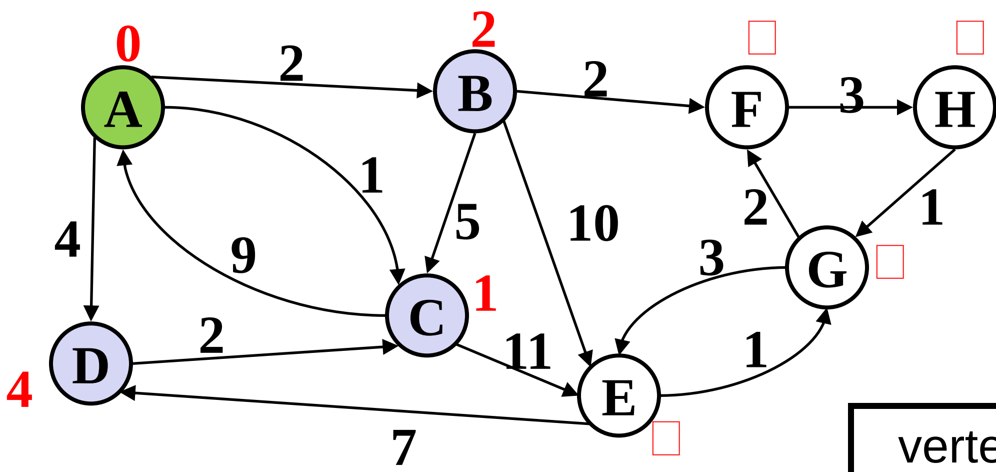
Example #1



vertex	known?	cost	path
A		0	
B		??	
C		??	
D		??	
E		??	
F		??	
G		??	
H		??	

<u>Order Added to Known Set:</u>

Example #1

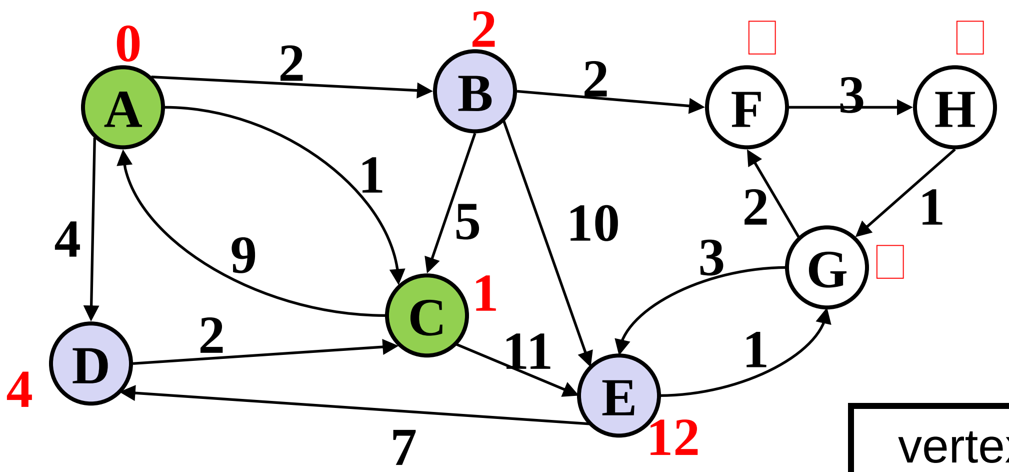


vertex	known?	cost	path
A	Y	0	
B		≤ 2	A
C		≤ 1	A
D		≤ 4	A
E		??	
F		??	
G		??	
H		??	

<u>Order Added to Known Set:</u>

A

Example #1

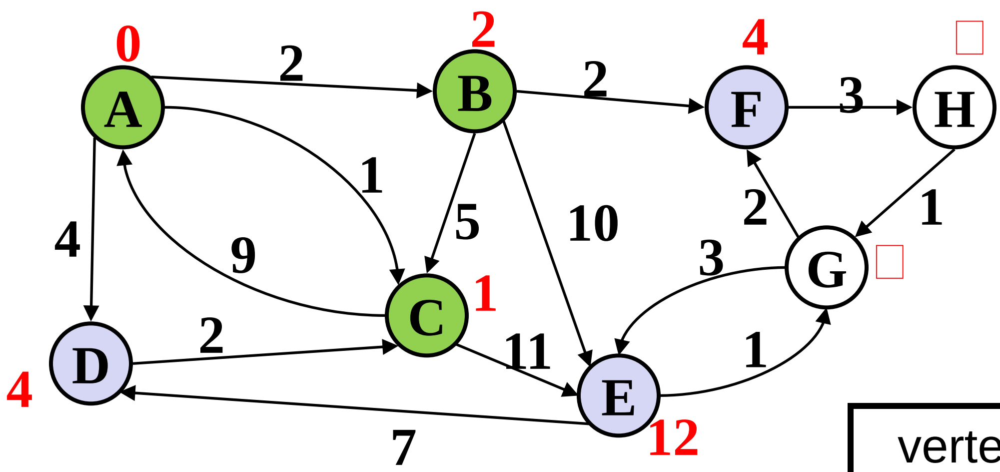


vertex	known?	cost	path
A	Y	0	
B		≤ 2	A
C	Y	1	A
D		≤ 4	A
E		≤ 12	C
F		??	
G		??	
H		??	

<u>Order Added to Known Set:</u>

A, C

Example #1

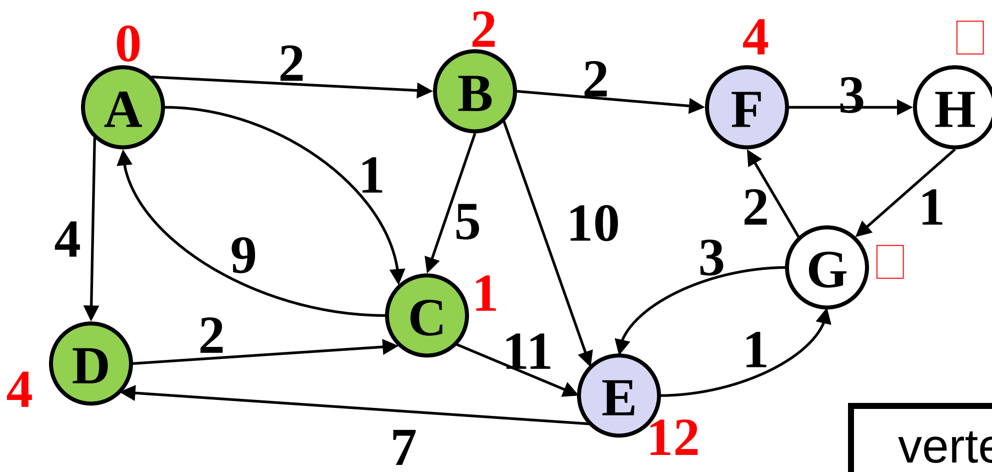


vertex	known?	cost	path
A	Y	0	
B	Y	2	A
C	Y	1	A
D		≤ 4	A
E		≤ 12	C
F		≤ 4	B
G		??	
H		??	

<u>Order Added to Known Set:</u>

A, C, B

Example #1

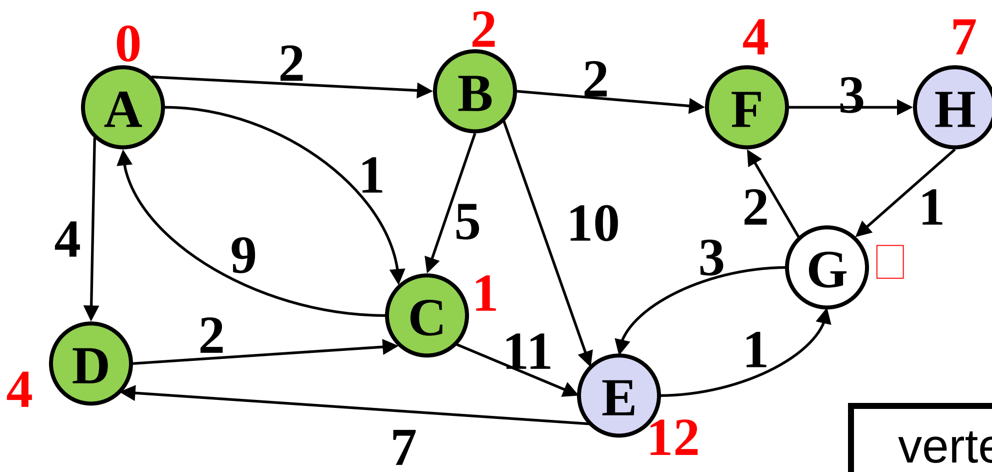


vertex	known?	cost	path
A	Y	0	
B	Y	2	A
C	Y	1	A
D	Y	4	A
E		≤ 12	C
F		≤ 4	B
G		??	
H		??	

<u>Order Added to Known Set:</u>

A, C, B, D

Example #1

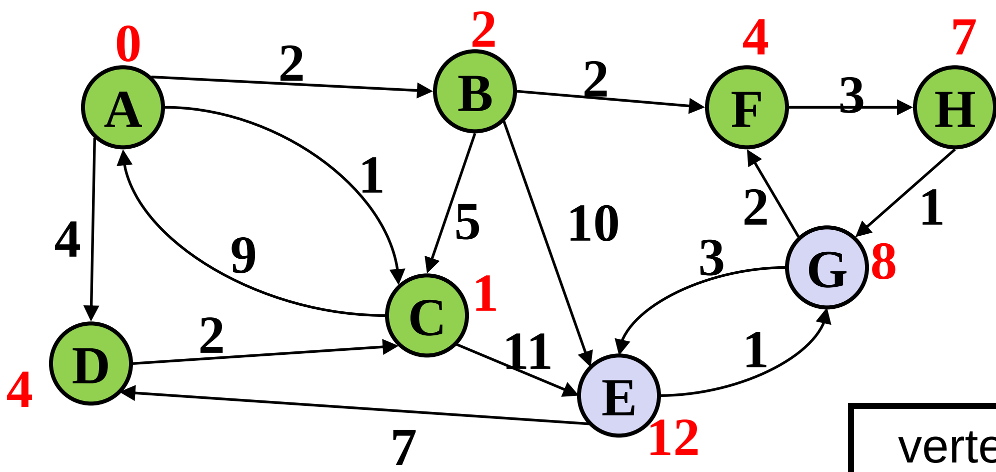


vertex	known?	cost	path
A	Y	0	
B	Y	2	A
C	Y	1	A
D	Y	4	A
E		≤ 12	C
F	Y	4	B
G		??	
H		≤ 7	F

<u>Order Added to Known Set:</u>

A, C, B, D, F

Example #1

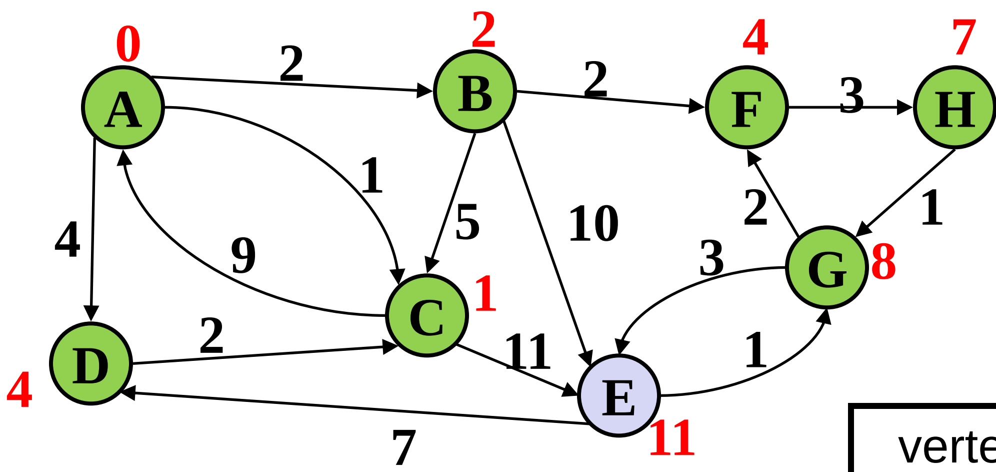


vertex	known?	cost	path
A	Y	0	
B	Y	2	A
C	Y	1	A
D	Y	4	A
E		≤ 12	C
F	Y	4	B
G		≤ 8	H
H	Y	7	F

<u>Order Added to Known Set:</u>

A, C, B, D, F, H

Example #1

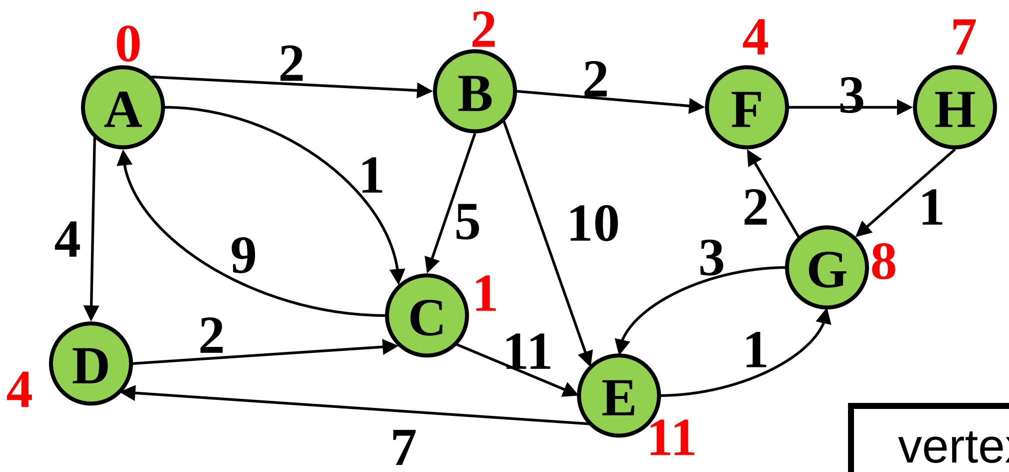


vertex	known?	cost	path
A	Y	0	
B	Y	2	A
C	Y	1	A
D	Y	4	A
E		≤ 11	G
F	Y	4	B
G	Y	8	H
H	Y	7	F

<u>Order Added to Known Set:</u>

A, C, B, D, F, H, G

Example #1



vertex	known?	cost	path
A	Y	0	
B	Y	2	A
C	Y	1	A
D	Y	4	A
E	Y	11	G
F	Y	4	B
G	Y	8	H
H	Y	7	F

<u>Order Added to Known Set:</u>

A, C, B, D, F, H, G, E

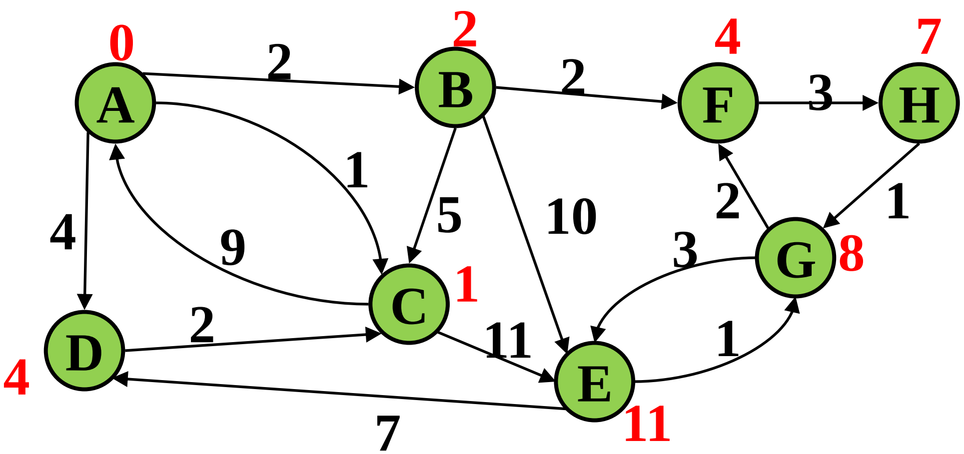
Important Features

- When a vertex is marked known,
  the cost of the shortest path to that node is known
  - The path is also known by following back-pointers

- While a vertex is still not known,
  another shorter path to it might still be found

Interpreting the Results

- Now that we're done, how do we get the path from, say, A to E?



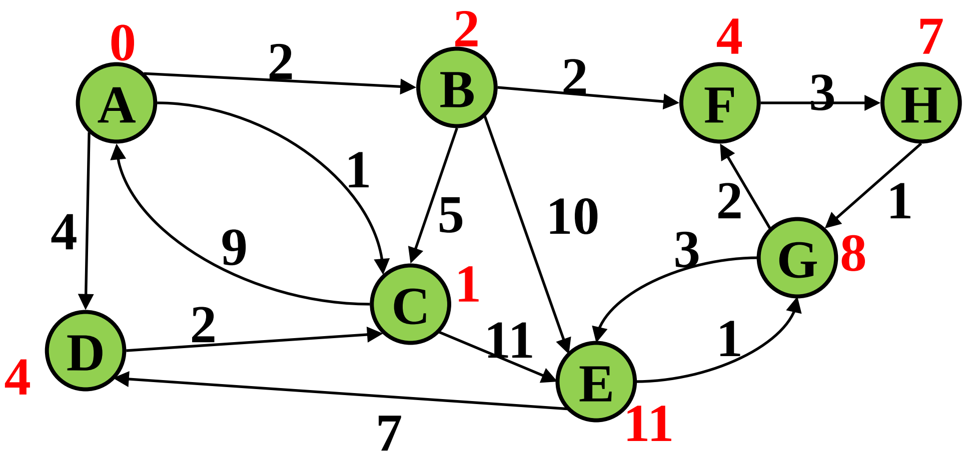
<u>Order Added to Known Set:</u>

A, C, B, D, F, H, G, E

vertex	known?	cost	path
A	Y	0	
B	Y	2	A
C	Y	1	A
D	Y	4	A
E	Y	11	G
F	Y	4	B
G	Y	8	H
H	Y	7	F

Stopping Short

- How would this have worked differently if we were only interested in:
  - the path from A to G?
  - the path from A to E?

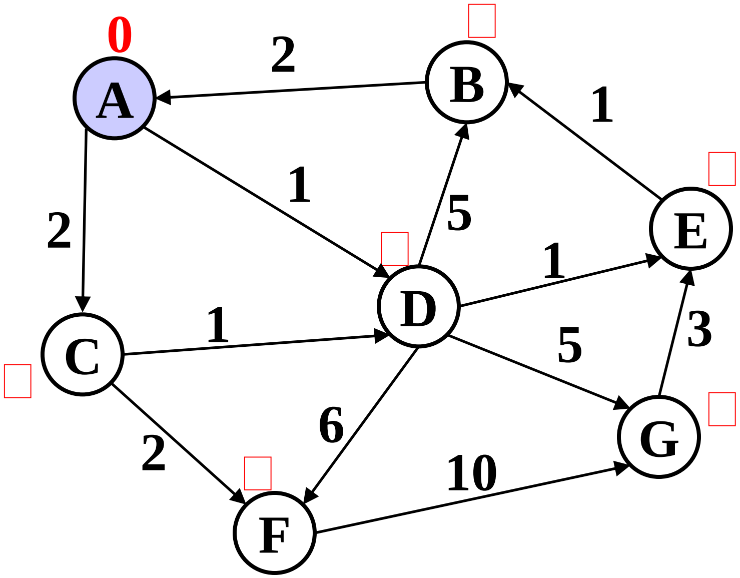


vertex	known?	cost	path
A	Y	0	
B	Y	2	A
C	Y	1	A
D	Y	4	A
E	Y	11	G
F	Y	4	B
G	Y	8	H
H	Y	7	F

<u>Order Added to Known Set:</u>

A, C, B, D, F, H, G, E

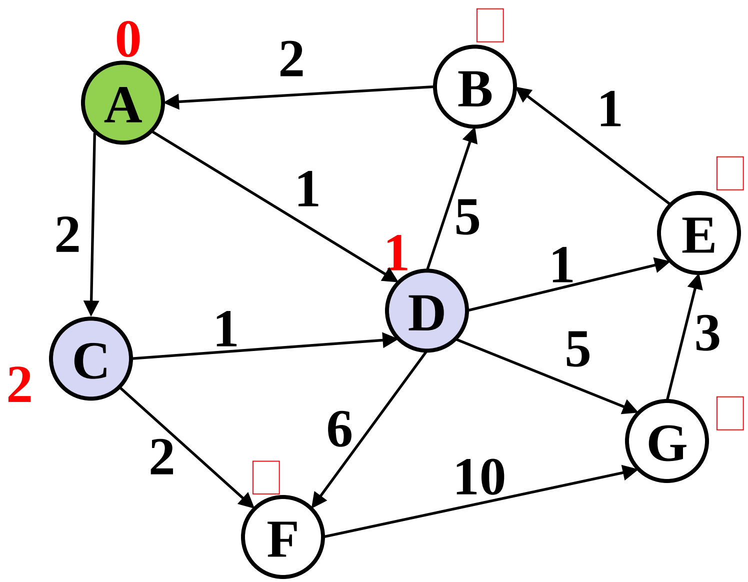
Example #2



Order Added to Known Set:

vertex	known?	cost	path
A		0	
B		??	
C		??	
D		??	
E		??	
F		??	
G		??	

Example #2

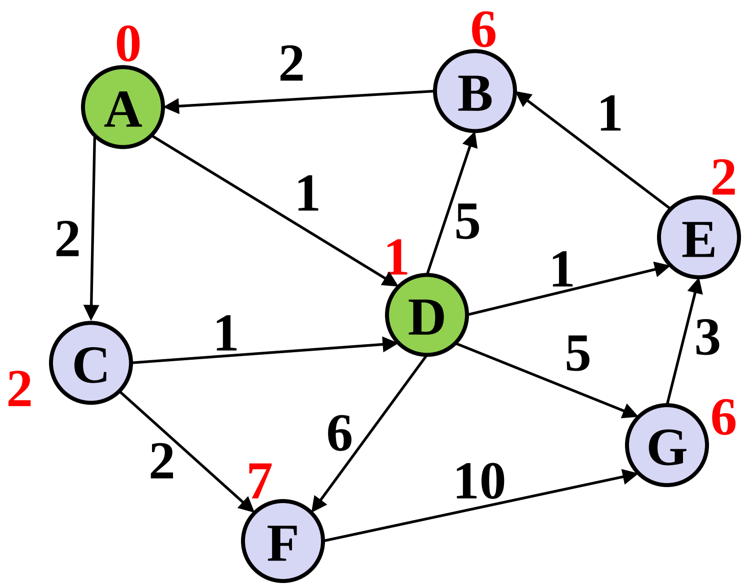


vertex	known?	cost	path
A	Y	0	
B		??	
C		≤ 2	A
D		≤ 1	A
E		??	
F		??	
G		??	

<u>Order Added to Known Set:</u>

A

Example #2

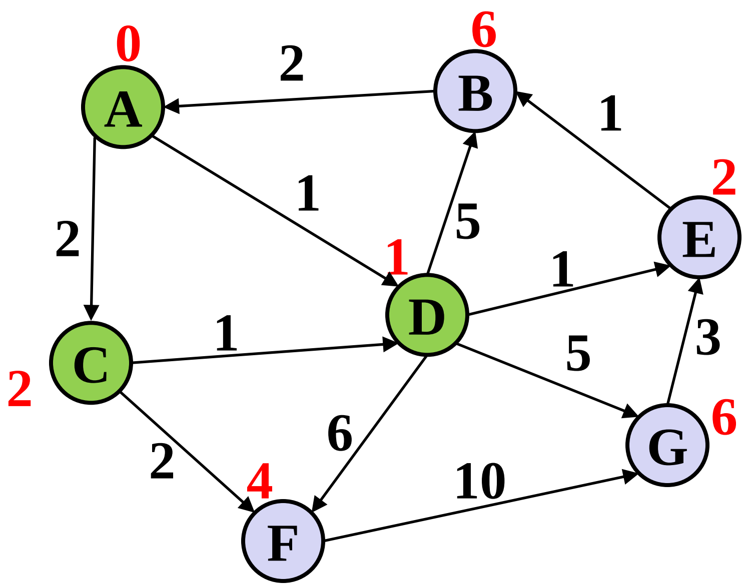


vertex	known?	cost	path
A	Y	0	
B		≤ 6	D
C		≤ 2	A
D	Y	1	A
E		≤ 2	D
F		≤ 7	D
G		≤ 6	D

<u>Order Added to Known Set:</u>

A, D

Example #2

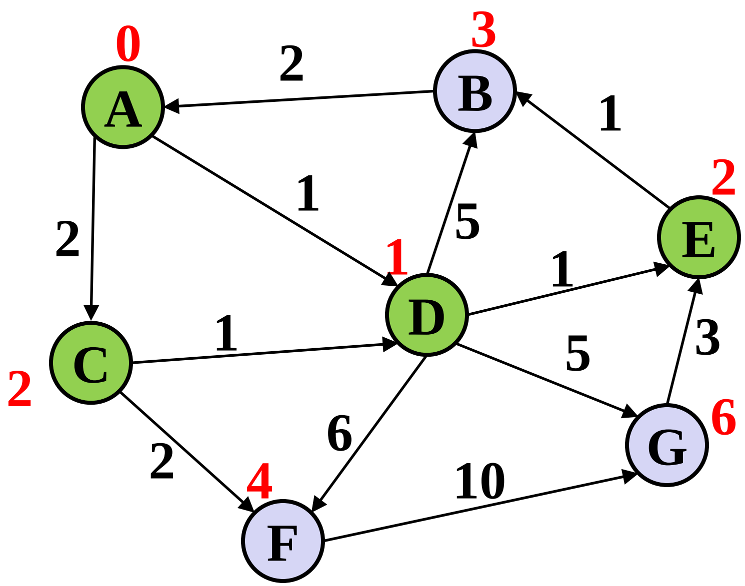


vertex	known?	cost	path
A	Y	0	
B		≤ 6	D
C	Y	2	A
D	Y	1	A
E		≤ 2	D
F		≤ 4	C
G		≤ 6	D

<u>Order Added to Known Set:</u>

A, D, C

Example #2

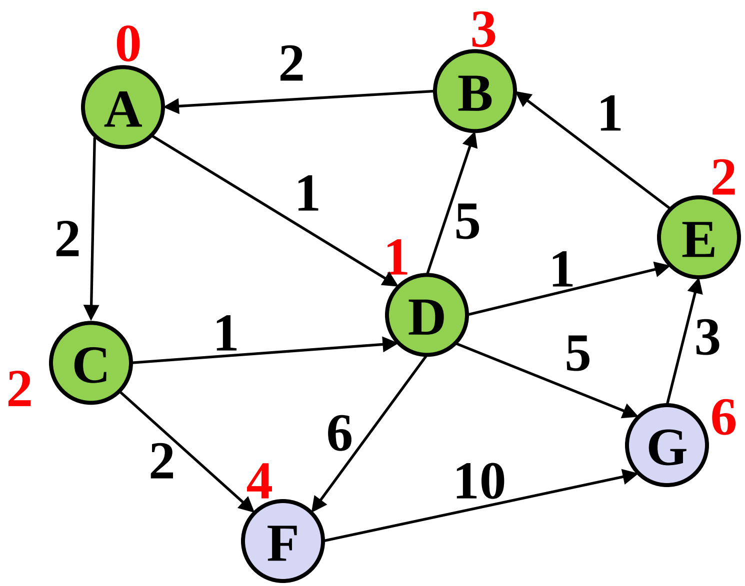


vertex	known?	cost	path
A	Y	0	
B		≤ 3	E
C	Y	2	A
D	Y	1	A
E	Y	2	D
F		≤ 4	C
G		≤ 6	D

<u>Order Added to Known Set:</u>

A, D, C, E

Example #2

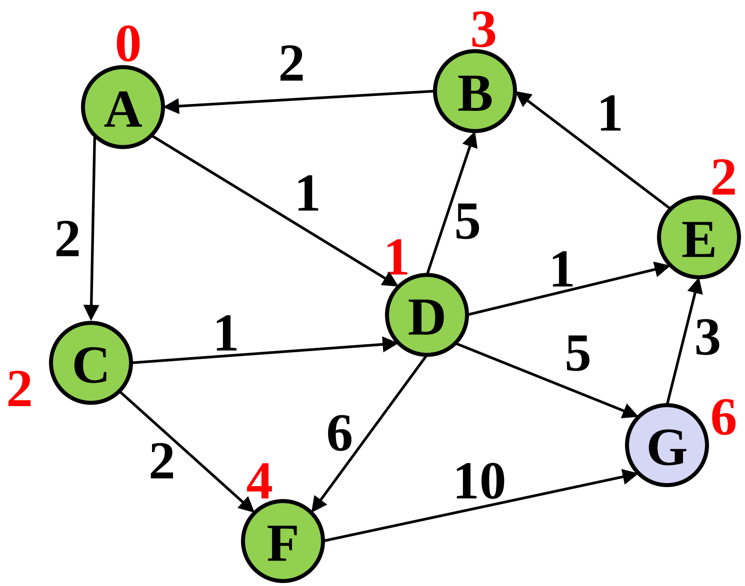


vertex	known?	cost	path
A	Y	0	
B	Y	3	E
C	Y	2	A
D	Y	1	A
E	Y	2	D
F		≤ 4	C
G		≤ 6	D

Order Added to Known Set:

A, D, C, E, B

Example #2

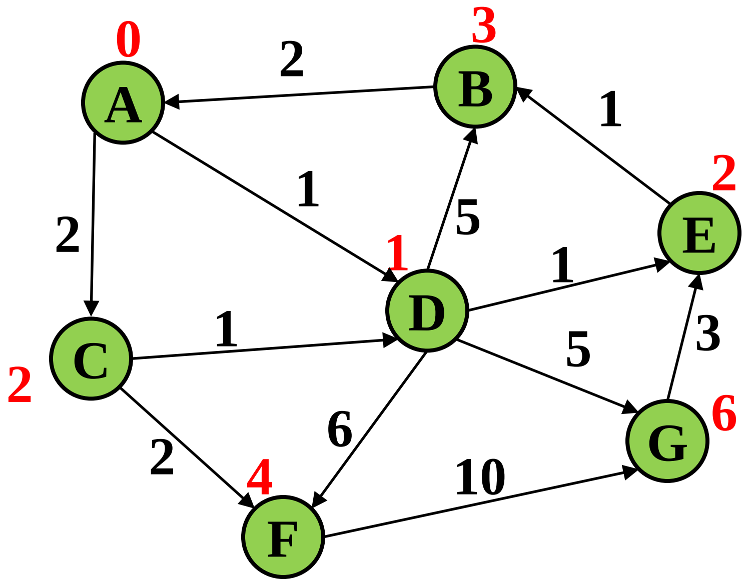


vertex	known?	cost	path
A	Y	0	
B	Y	3	E
C	Y	2	A
D	Y	1	A
E	Y	2	D
F	Y	4	C
G		≤ 6	D

<u>Order Added to Known Set:</u>

A, D, C, E, B, F

Example #2

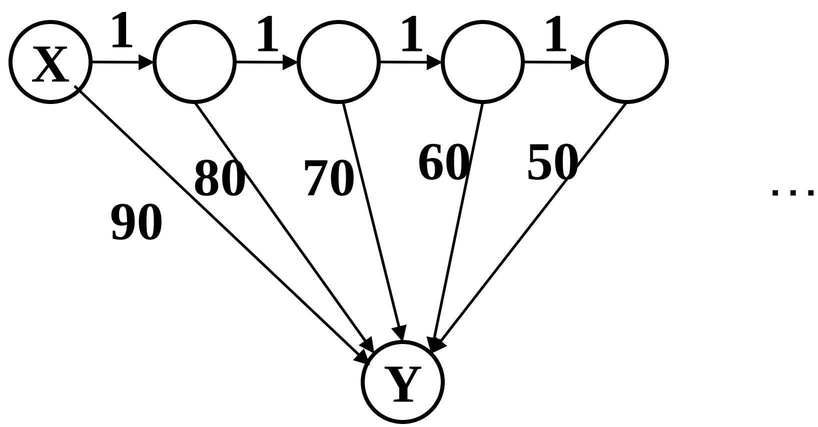


Order Added to Known Set:

A, D, C, E, B, F, G

vertex	known?	cost	path
A	Y	0	
B	Y	3	E
C	Y	2	A
D	Y	1	A
E	Y	2	D
F	Y	4	C
G	Y	6	D

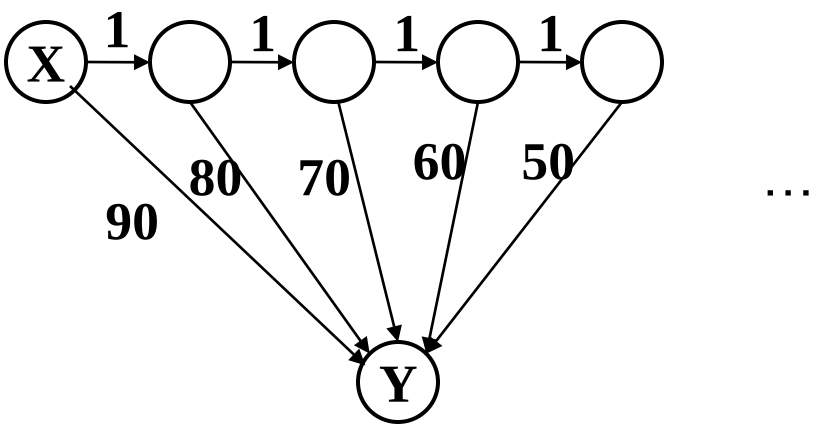
Example #3



How will the best-cost-so-far for Y proceed?

Is this expensive?

Example #3



How will the best-cost-so-far for Y proceed? 90, 81, 72, 63, 54, …

Is this expensive? No, each *edge* is processed only once

A Greedy Algorithm

- Dijkstra's algorithm
  - For single-source shortest paths in a weighted graph (directed or undirected) with no negative-weight edges

  - An example of a *greedy algorithm*:
    - At each step, irrevocably does what seems best at that step
      - once a vertex is in the known set,
        does not go back and readjust its decision
    - Locally optimal
      - does not always mean globally optimal

Where are We?

- Have described Dijkstra’s algorithm
  - For single-source shortest paths in a weighted graph (directed or undirected) with no negative-weight edges

- What should we do after learning an algorithm?
  - Prove it is correct
    - Not obvious!
    - We will sketch the key ideas
  - Analyze its efficiency
    - Will do better by using a data structure we learned earlier!

Correctness: Intuition

Rough intuition:

All the “known” vertices have the correct shortest path

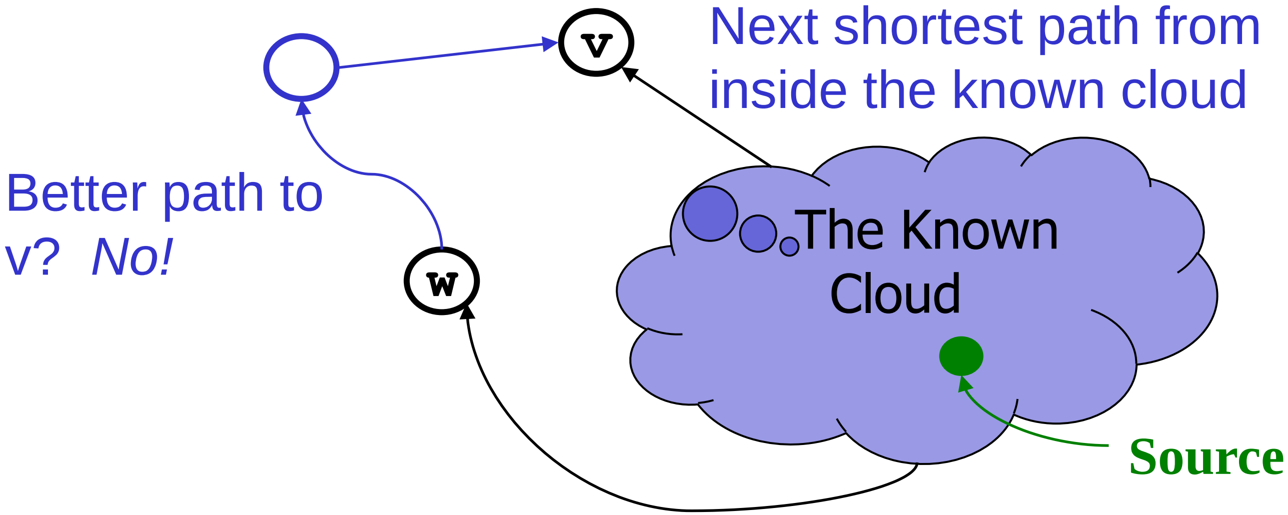
- True initially: shortest path to start node has cost 0
- If it stays true every time we mark a node “known”,
  then by induction this holds and eventually everything is “known”

Key fact we need:

When we mark a vertex “known” we won’t discover a shorter path later!

- This holds only because Dijkstra’s algorithm
  picks the node with the next shortest path-so-far
- The proof is by contradiction…

Correctness: The Cloud (Rough Sketch)



Suppose **v** is the next node to be marked known ("added to the cloud")

- The best-known path to **v** must have only nodes "in the cloud"
  - We have selected it, and we only know about paths through the cloud to a node at the edge of the cloud
- Assume the actual shortest path to **v** is different
  - It is not entirely within the cloud, or else we would know about it
    - So it must use non-cloud nodes
  - Let **w** be the *first* non-cloud node on this path.
  - The part of the path up to **w** is already known and must be shorter than the best-known path to **v**. So **v** would not have been picked. Contradiction.

Efficiency, First Approach

Use pseudocode to determine asymptotic run-time

- Notice each edge is processed only once

```
dijkstra(Graph G, Node start) {
  for each node: x.cost=infinity, x.known=false
  start.cost = 0
  while(not all nodes are known) {
    b = find unknown node with smallest cost
    b.known = true
    for each edge (b,a) in G
     if(!a.known)
       if(b.cost + weight((b,a)) < a.cost){
         a.cost = b.cost + weight((b,a))
         a.path = b
       }
}
```

Efficiency, First Approach

Use pseudocode to determine asymptotic run-time

- Notice each edge is processed only once

```
dijkstra(Graph G, Node start) {
  for each node: x.cost=infinity, x.known=false
  start.cost = 0
  while(not all nodes are known) {
    b = find unknown node with smallest cost
    b.known = true
    for each edge (b,a) in G
     if(!a.known)
       if(b.cost + weight((b,a)) < a.cost){
         a.cost = b.cost + weight((b,a))
         a.path = b
       }
}
```

$O(|V|)$

$O(|V|^2)$

$O(|E|)$

$O(|V|^2)$

Improving Asymptotic Running Time

- So far: $O(|V|^2)$
- We had a similar “problem” with topological sort being $O(|V|^2)$ due to each iteration looking for the node to process next
  - We solved it with a queue of zero-degree nodes
  - But here we need the lowest-cost node and costs can change as we process edges
- Solution?

Improving Asymptotic Running Time

- So far: $O(|V|^2)$
- We had a similar "problem" with topological sort being $O(|V|^2)$ due to each iteration looking for the node to process next
  - We solved it with a queue of zero-degree nodes
  - But here we need the lowest-cost node and costs can change as we process edges
- Solution?
  - A priority queue holding all unknown nodes, sorted by cost
  - But must support **`decreaseKey`** operation
    - Must maintain a reference from each node to its position in the priority queue
    - Conceptually simple, but can be a pain to code up

Efficiency, Second Approach

Use pseudocode to determine asymptotic run-time

```
dijkstra(Graph G, Node start) {
  for each node: x.cost=infinity, x.known=false
  start.cost = 0
  build-heap with all nodes
  while(heap is not empty) {
    b = deleteMin()
    b.known = true
    for each edge (b,a) in G
     if(!a.known)
      if(b.cost + weight((b,a)) < a.cost){
        decreaseKey(a,"new cost – old cost")
        a.path = b
      }
}
```

Efficiency, Second Approach

Use pseudocode to determine asymptotic run-time

```
dijkstra(Graph G, Node start) {
  for each node: x.cost=infinity, x.known=false
  start.cost = 0
  build-heap with all nodes
  while(heap is not empty) {
    b = deleteMin()
    b.known = true
    for each edge (b,a) in G
     if(!a.known)
      if(b.cost + weight((b,a)) < a.cost){
        decreaseKey(a,"new cost – old cost")
        a.path = b
      }
}
```

O(|V|)

O(|V|log|V|)

O(|E|log|V|)

O(|V|log|V|+|E|log|V|)

Dense vs. Sparse Again

- First approach: $O(|V|^2)$
- Second approach: $O(|V|\log|V|+|E|\log|V|)$
- So which is better?
  - Sparse: $O(|V|\log|V|+|E|\log|V|)$ (if $|E| > |V|$, then $O(|E|\log|V|)$)
  - Dense: $O(|V|^2)$
- But, remember these are worst-case and asymptotic
  - Priority queue might have slightly worse constant factors
  - On the other hand, for “normal graphs”, we might rarely call **`decreaseKey`** (or not percolate far), making $|E|\log|V|$ more like $|E|$

All-Pairs Shortest Path

- Find the shortest path between all pairs of vertices in the graph

- How?

Dynamic Programming

Algorithmic technique that systematically records the answers to sub-problems in a table and re-uses those recorded results (rather than re-computing them).

Simple Example: Calculating the Nth Fibonacci number.

Fib(N) = Fib(N-1) + Fib(N-2)

Recursion would be insanely expensive,
but it is cheap if you already know results of prior computations

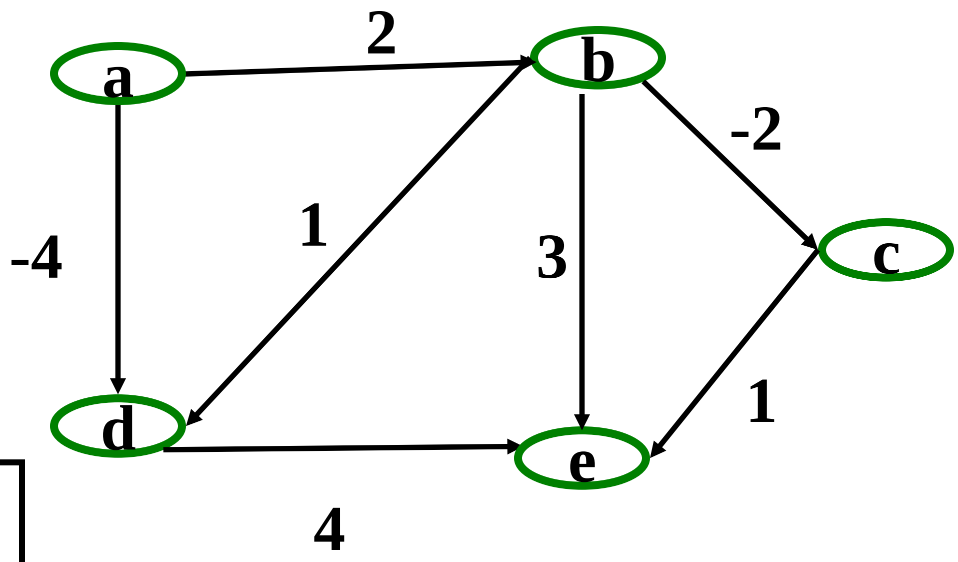
Floyd-Warshall

```
for (int k = 1; k =< V; k++)
 for (int i = 1; i =< V; i++)
  for (int j = 1; j =< V; j++)
   if ( ( M[i][k]+ M[k][j] ) < M[i][j] )
     M[i][j] =  M[i][k]+ M[k][j]
```

Invariant:

After the kth iteration, for all pairs of vertices the matrix includes the shortest path containing only vertices 1..k as intermediate vertices

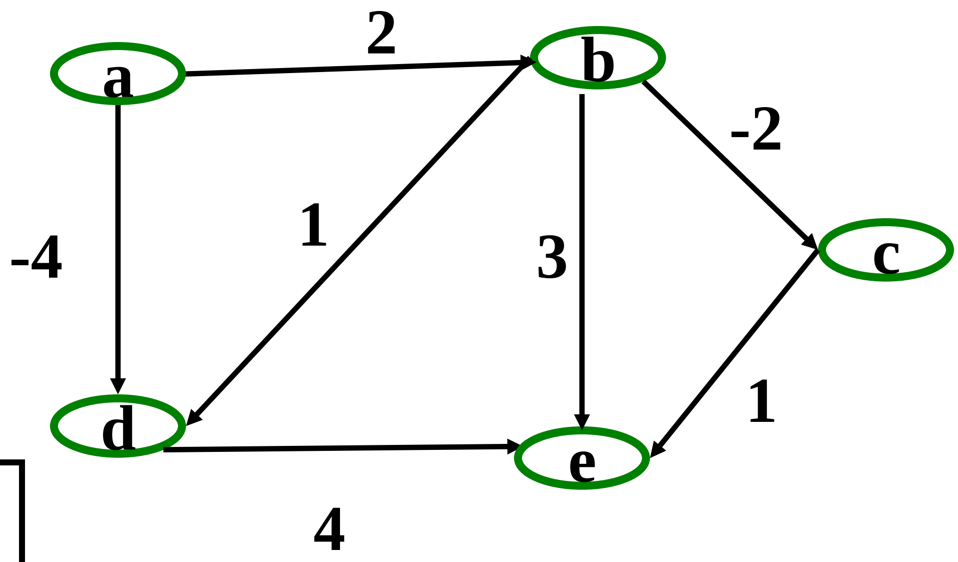
Initial state of the matrix:



	a	b	c	d	e
a	0	2	-	-4	-
b	-	0	-2	1	3
c	-	-	0	-	1
d	-	-	-	0	4
e	-	-	-	-	0

$$M[i][j] = \min(M[i][j], M[i][k] + M[k][j])$$

Initial state of the matrix:

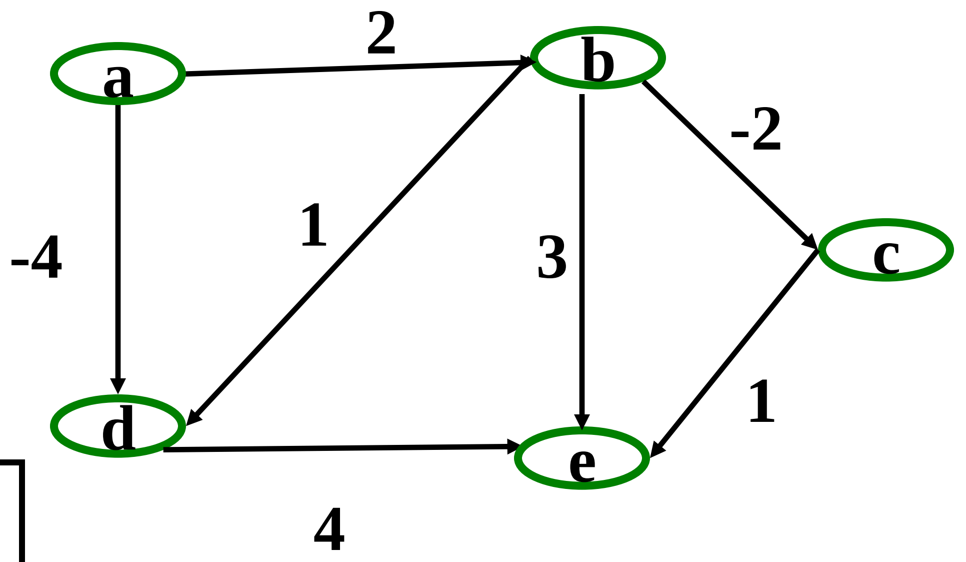


	a	b	c	d	e
a	0	2	-	-4	-
b	-	0	-2	1	3
c	-	-	0	-	1
d	-	-	-	0	4
e	-	-	-	-	0

k = 1

$$M[i][j] = \min(M[i][j], M[i][k] + M[k][j])$$

Initial state of the matrix:

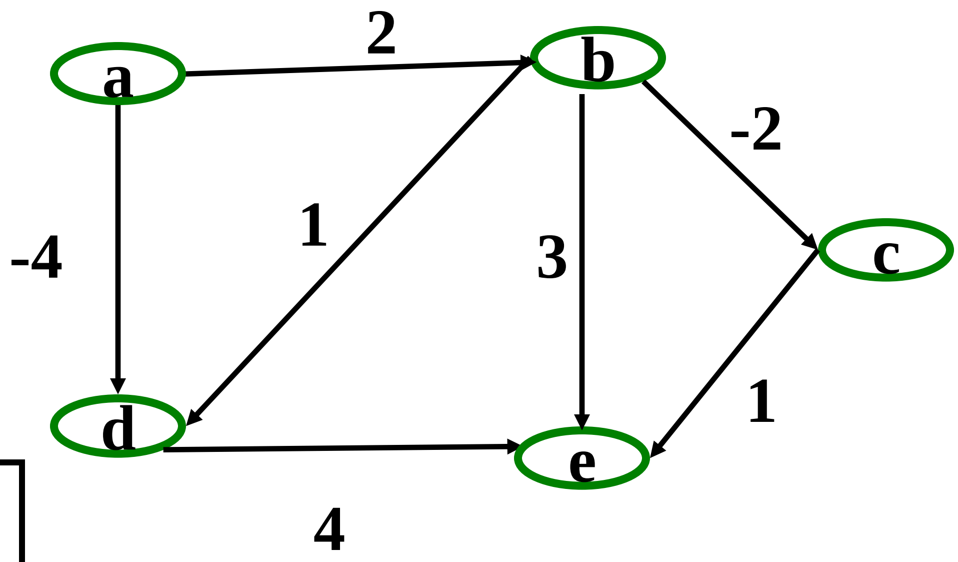


	a	b	c	d	e
a	0	2	-	-4	-
b	-	0	-2	1	3
c	-	-	0	-	1
d	-	-	-	0	4
e	-	-	-	-	0

k = 2

$$M[i][j] = \color{red}{\min}(M[i][j], M[i][\color{green}{k}] + M[\color{green}{k}][j])$$

Initial state of the matrix:

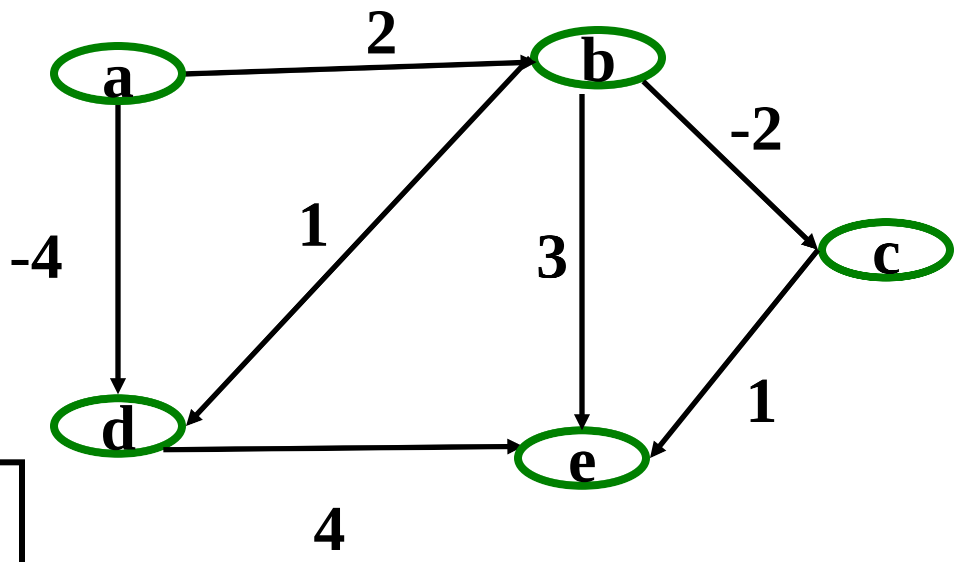


	a	b	c	d	e
a	0	2	0	-4	5
b	-	0	-2	1	3
c	-	-	0	-	1
d	-	-	-	0	4
e	-	-	-	-	0

k = 2

$$M[i][j] = \min(M[i][j], M[i][k] + M[k][j])$$

Initial state of the matrix:

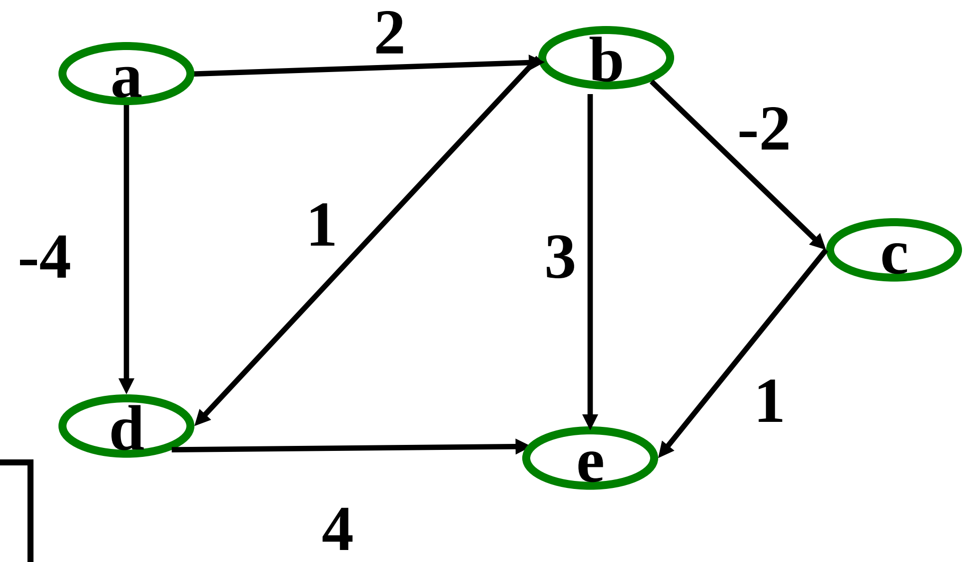


	a	b	c	d	e
a	0	2	0	-4	5
b	-	0	-2	1	3
c	-	-	0	-	1
d	-	-	-	0	4
e	-	-	-	-	0

k = 3

$$M[i][j] = \min(M[i][j], M[i][k] + M[k][j])$$

Initial state of the matrix:

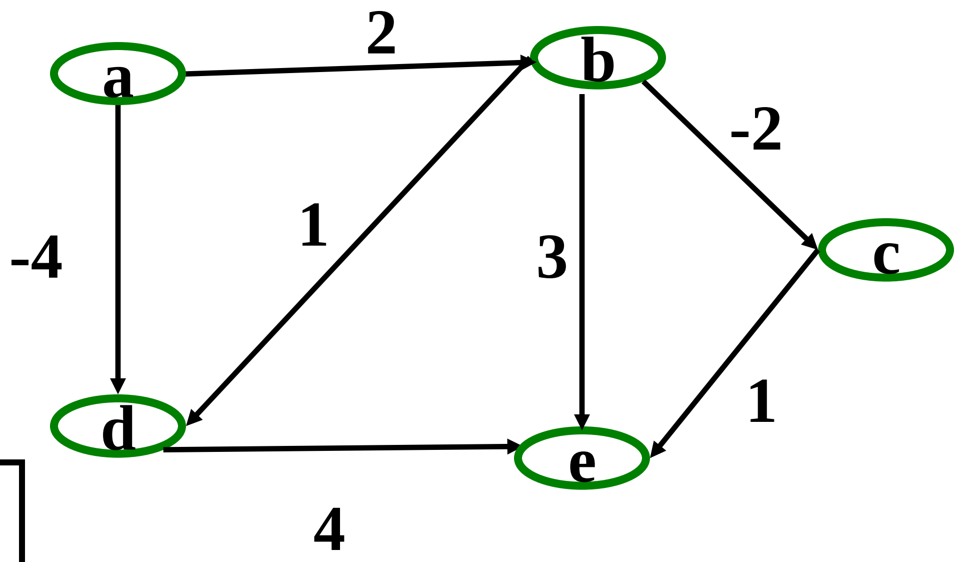


	a	b	c	d	e
a	0	2	0	-4	1
b	-	0	-2	1	-1
c	-	-	0	-	1
d	-	-	-	0	4
e	-	-	-	-	0

k = 3

M[i][j] = min(M[i][j], M[i][k]+ M[k][j])

Initial state of the matrix:

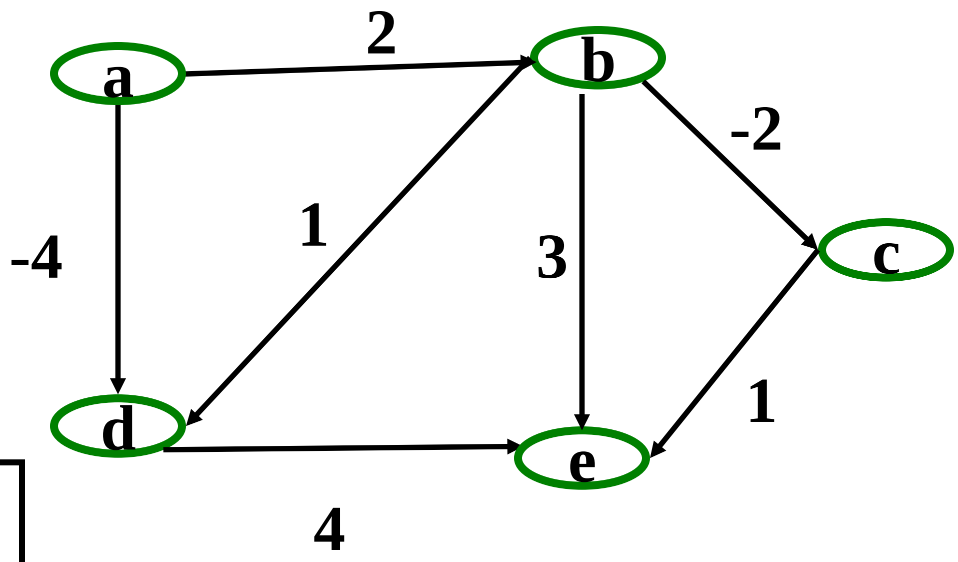


	a	b	c	d	e
a	0	2	0	-4	1
b	-	0	-2	1	-1
c	-	-	0	-	1
d	-	-	-	0	4
e	-	-	-	-	0

k = 4

$$M[i][j] = \min(M[i][j],\ M[i][k] + M[k][j])$$

Initial state of the matrix:

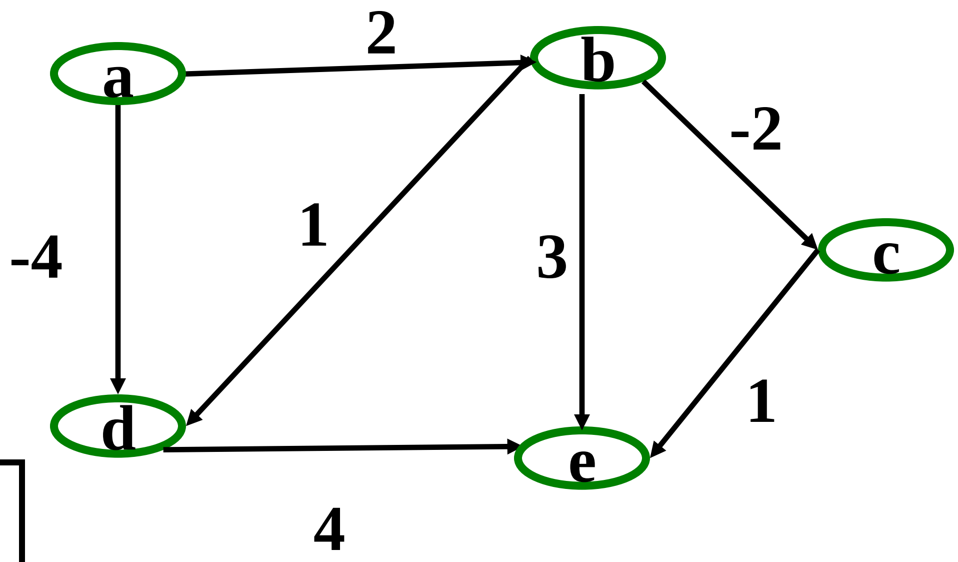


	a	b	c	d	e
a	0	2	0	-4	0
b	-	0	-2	1	-1
c	-	-	0	-	1
d	-	-	-	0	4
e	-	-	-	-	0

k = 4

$$M[i][j] = \min(M[i][j], M[i][k] + M[k][j])$$

Initial state of the matrix:



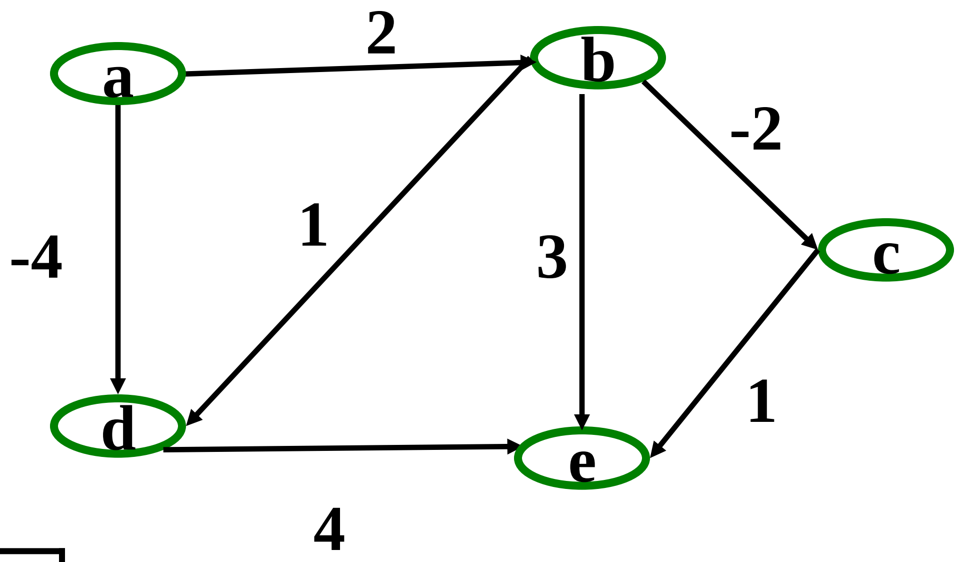
	a	b	c	d	e
a	0	2	0	-4	0
b	-	0	-2	1	-1
c	-	-	0	-	1
d	-	-	-	0	4
e	-	-	-	-	0

k = 5

M[i][j] = min(M[i][j], M[i][k]+ M[k][j])

Floyd-Warshall

All-Pairs Shortest Path



	a	b	c	d	e
a	0	2	0	-4	0
b	-	0	-2	1	-1
c	-	-	0	-	1
d	-	-	-	0	4
e	-	-	-	-	0

Final Matrix Contents

What Comes Next?

In the logical course progression, we would next study

1. Minimum spanning trees

But to align lectures with projects and homeworks, instead we will

- Start parallelism and concurrency
- Come back to graphs at the end of the course

Note toward the future:

- We cannot do all of graphs last because of the CSE312 co-requisite (needed for study of NP)