CSE 421
Introduction to Algorithms

Lecture 6: More Greedy Algorithms

Last time: Greedy Algorithms

Hard to define exactly but can give general properties

- Solution is built in small steps
- Decisions on how to build the solution are made to maximize some criterion without looking to the future
  - Want the 'best' current partial solution as if the current step were the last step

May be more than one greedy algorithm using different criteria to solve a given problem

- Not obvious which criteria will actually work

Greedy Analysis Strategies

Greedy algorithm stays ahead: Show that after each step of the greedy algorithm, its solution is at least as good as any other algorithm's

- Example: Interval Scheduling analysis

Structural: Discover a simple "structural" bound asserting that every possible solution must have a certain value. Then show that your algorithm always achieves this bound.

- Example: Interval Partitioning analysis

Exchange argument: Gradually transform any solution to the one found by the greedy algorithm without hurting its quality.

Scheduling to Minimize Lateness

Scheduling to minimize lateness:

- Single resource as in interval scheduling but, instead of start and finish times, request $\boldsymbol{i}$ has
  - Time requirement $\boldsymbol{t_i}$ which must be scheduled in a contiguous block
  - Target deadline $\boldsymbol{d_i}$ by which time the request would like to be finished
- Overall start time $\boldsymbol{s}$ for all jobs

Requests are scheduled by the algorithm into time intervals $[\boldsymbol{s_i}, \boldsymbol{f_i}]$ s.t. $\boldsymbol{t_i = f_i - s_i}$

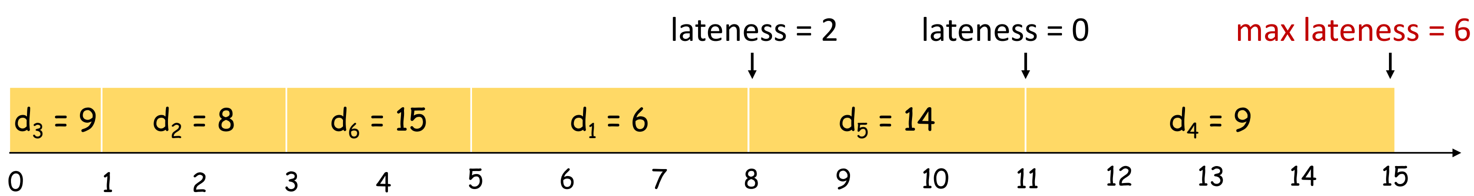
- Lateness of schedule for request $\boldsymbol{i}$ is
  - If $\boldsymbol{f_i > d_i}$ then request $\boldsymbol{i}$ is late by $\boldsymbol{L_i = f_i - d_i}$; otherwise its lateness $\boldsymbol{L_i = 0}$
- Maximum lateness $\boldsymbol{L} = \max_i \boldsymbol{Li}$

Goal: Find a schedule for **all** requests (values of $\boldsymbol{s_i}$ and $\boldsymbol{f_i}$ for each request $\boldsymbol{i}$) to minimize the maximum lateness, $\boldsymbol{L}$.

Scheduling to Minimizing Lateness

- Example:

	1	2	3	4	5	6
t_j	3	2	1	4	3	2
d_j	6	8	9	9	14	15



Minimizing Lateness: Greedy Algorithms

Greedy template: Consider jobs in some order.

[Shortest processing time first] Consider jobs in ascending order of processing time t_j.

[Earliest deadline first] Consider jobs in ascending order of deadline d_j.

[Smallest slack] Consider jobs in ascending order of slack $d_j - t_j$.

Minimizing Lateness: Greedy Algorithms

Greedy template: Consider jobs in some order.

[Shortest processing time first] Consider jobs in ascending order of processing time t_j.

	1	2
t_j	1	10
d_j	100	10

counterexample

Will schedule 1 (length 1) before 2 (length 10).
2 can only be scheduled at time 1
1 will finish at time 11 >10. Lateness 1.
Lateness 0 possible If 1 goes last.

[Smallest slack] Consider jobs in ascending order of slack $d_j - t_j$.

	1	2
t_j	1	10
d_j	2	10

counterexample

Will schedule 2 (slack 0) before 1 (slack 1).
1 can only be scheduled at time 10
1 will finish at time 11 >10. Lateness 9.
Lateness 1 possible if 1 goes first.

Minimizing Lateness: Greedy Algorithms

Greedy template: Consider jobs in some order.

[Earliest deadline first] Consider jobs in ascending order of deadline d_j.

Greedy Algorithm: Earliest Deadline First

Consider requests in increasing order of deadlines

Schedule the request with the earliest deadline as soon as the resource is available

Minimizing Lateness: Greedy EDF Algorithm

- Greedy Earliest Deadline First (EDF).

Sort deadlines in increasing order $(d_1 \leq d_2 \leq \cdots \leq d_n)$
$f \leftarrow s$
for $i \leftarrow 1$ to n {
    $s_i \leftarrow f$
    $f_i \leftarrow s_i + t_i$
    $f \leftarrow f_i$
}

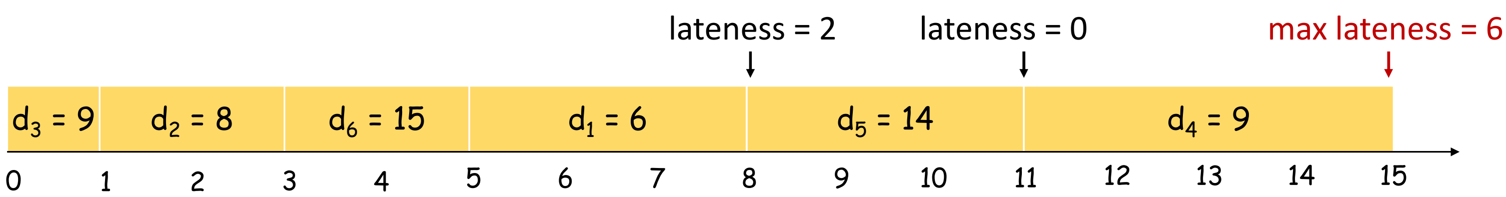
More on Monday!

Scheduling to Minimizing Lateness

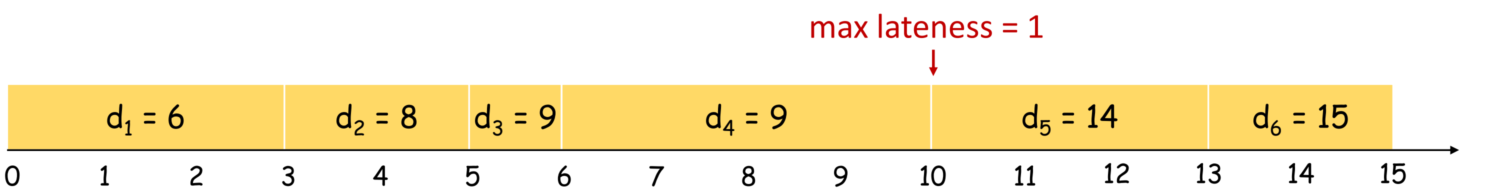
- Example:

	1	2	3	4	5	6
t_j	3	2	1	4	3	2
d_j	6	8	9	9	14	15

Original Schedule



EDF Schedule



Proof for Greedy EDF Algorithm: Exchange Argument

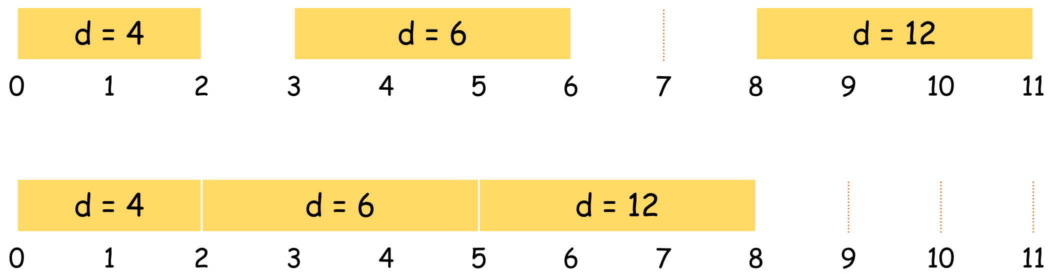
Show that if there is another schedule **O** (think optimal schedule) then we can gradually change **O** so that…

- at each step the maximum lateness in **O** never gets worse
- it eventually becomes the same cost as **A**

This means that **A** is at least as good as **O**, so **A** is also optimal!

Minimizing Lateness: No Idle Time

Observation: There exists an optimal schedule with no idle time

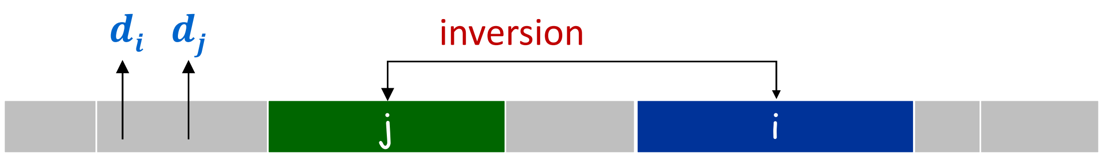


At least as good

Observation: The greedy EDF schedule has no idle time.

Minimizing Lateness: Inversions

Defn: An inversion in schedule S is a pair of jobs i and j such that $d_i < d_j$ but j is scheduled before i.



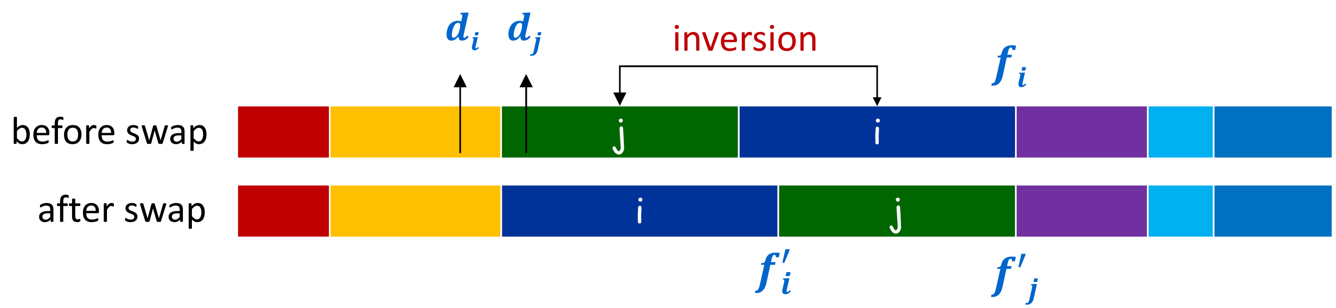
Observation: Greedy EDF schedule has no inversions.

Observation: If schedule S (with no idle time) has an inversion it has two adjacent jobs that are inverted

- Any job in between would be inverted w.r.t. one of the two ends

Minimizing Lateness: Inversions

Defn: An inversion in schedule S is a pair of jobs i and j such that $d_i < d_j$ but j is scheduled before i.

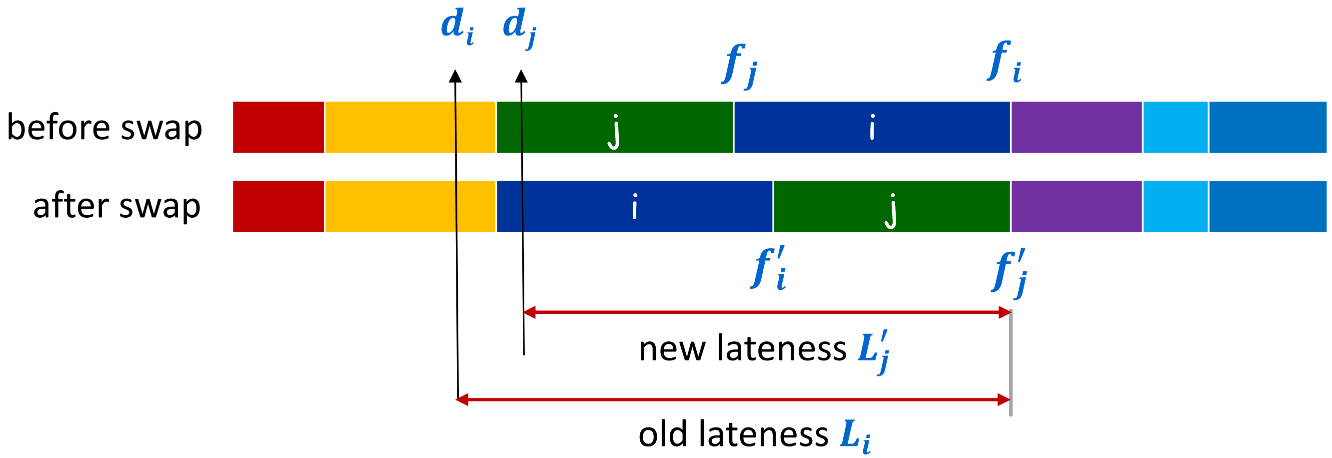


Claim: Swapping two adjacent, inverted jobs

- reduces the # of inversions by 1
- does not increase the max lateness.

Minimizing Lateness: Inversions

Defn: An inversion in schedule S is a pair of jobs i and j such that $d_i < d_j$ but j is scheduled before i.



Claim: Maximum lateness does not increase

Optimal schedules and inversions

Claim: There is an optimal schedule with no idle time and no inversions

Proof:

By previous argument there is an optimal schedule **O** with no idle time

If **O** has an inversion then it has an **adjacent** pair of requests in its schedule that are inverted and can be swapped without increasing lateness

… we just need to show one more claim that eventually this swapping stops

Optimal schedules and inversions

Claim: Eventually these swaps will produce an optimal schedule with no inversions.

Proof:

Each swap decreases the # of inversions by $\mathbf{1}$

There are a bounded # of inversions possible in the worst case

- at most $\boldsymbol{n(n-1)/2}$ but we only care that this is finite.

The # of inversions can't be negative so this must stop. ■

Idleness and Inversions are the only issue

Claim: All schedules with no inversions and no idle time have the same maximum lateness.

Proof:

Schedules can differ only in how they order requests with equal deadlines

Consider all requests having some common deadline d.

- Maximum lateness of these jobs is based only on finish time of the last one …

and the set of these requests occupies the same time segment in both schedules.

$\Rightarrow$ The last of these requests finishes at the same time in any such schedule. $\blacksquare$

Earliest Deadline First is optimal

We know that

- There is an optimal schedule with no idle time or inversions
- All schedules with no idle time or inversions have the same maximum lateness
- EDF produces a schedule with no idle time or inversions

So …

- EDF produces an optimal schedule ∎

Greedy Analysis Strategies

Greedy algorithm stays ahead: Show that after each step of the greedy algorithm, its solution is at least as good as any other algorithm's

Structural: Discover a simple "structural" bound asserting that every possible solution must have a certain value. Then show that your algorithm always achieves this bound.

Exchange argument: Gradually transform any solution to the one found by the greedy algorithm without hurting its quality.

Single-source shortest paths

Given: an (un)directed graph $\boldsymbol{G = (V, E)}$ with each edge $\boldsymbol{e}$ having a non-negative weight $\boldsymbol{w(e)}$ and a vertex $\boldsymbol{s}$

Find: (length of) shortest paths from $\boldsymbol{s}$ to each vertex in $\boldsymbol{G}$

A Greedy Algorithm

Dijkstra's Algorithm:

- Maintain a set $\boldsymbol{S}$ of vertices whose shortest paths are known
  - initially $\boldsymbol{S} = \{\boldsymbol{s}\}$
- Maintaining current best lengths of paths that *only go through* $\boldsymbol{S}$ to each of the vertices in $\boldsymbol{G}$
  - path-lengths to elements of $\boldsymbol{S}$ will be right, to $\boldsymbol{V} \setminus \boldsymbol{S}$ they might not be right
- Repeatedly add vertex $\boldsymbol{v}$ to $\boldsymbol{S}$ that has the shortest path-length of any vertex in $\boldsymbol{V} \setminus \boldsymbol{S}$
  - update path lengths based on new paths through $\boldsymbol{v}$

Dijsktra's Algorithm

Dijkstra($\boldsymbol{G}$,$\boldsymbol{w}$,s)

$\boldsymbol{S} \leftarrow \{\boldsymbol{s}\}$

$\boldsymbol{d}[\boldsymbol{s}] \leftarrow \boldsymbol{0}$

while $\boldsymbol{S} \neq \boldsymbol{V}$ {

among all edges $\boldsymbol{e} = (\boldsymbol{u}, \boldsymbol{v})$ s.t. $\boldsymbol{v} \notin \boldsymbol{S}$ and $\boldsymbol{u} \in \boldsymbol{S}$ select* one with the minimum value of $\boldsymbol{d}[\boldsymbol{u}] + \boldsymbol{w}(\boldsymbol{e})$

$\boldsymbol{S} \leftarrow \boldsymbol{S} \cup \{\boldsymbol{v}\}$

$\boldsymbol{d}[\boldsymbol{v}] \leftarrow \boldsymbol{d}[\boldsymbol{u}] + \boldsymbol{w}(\boldsymbol{e})$

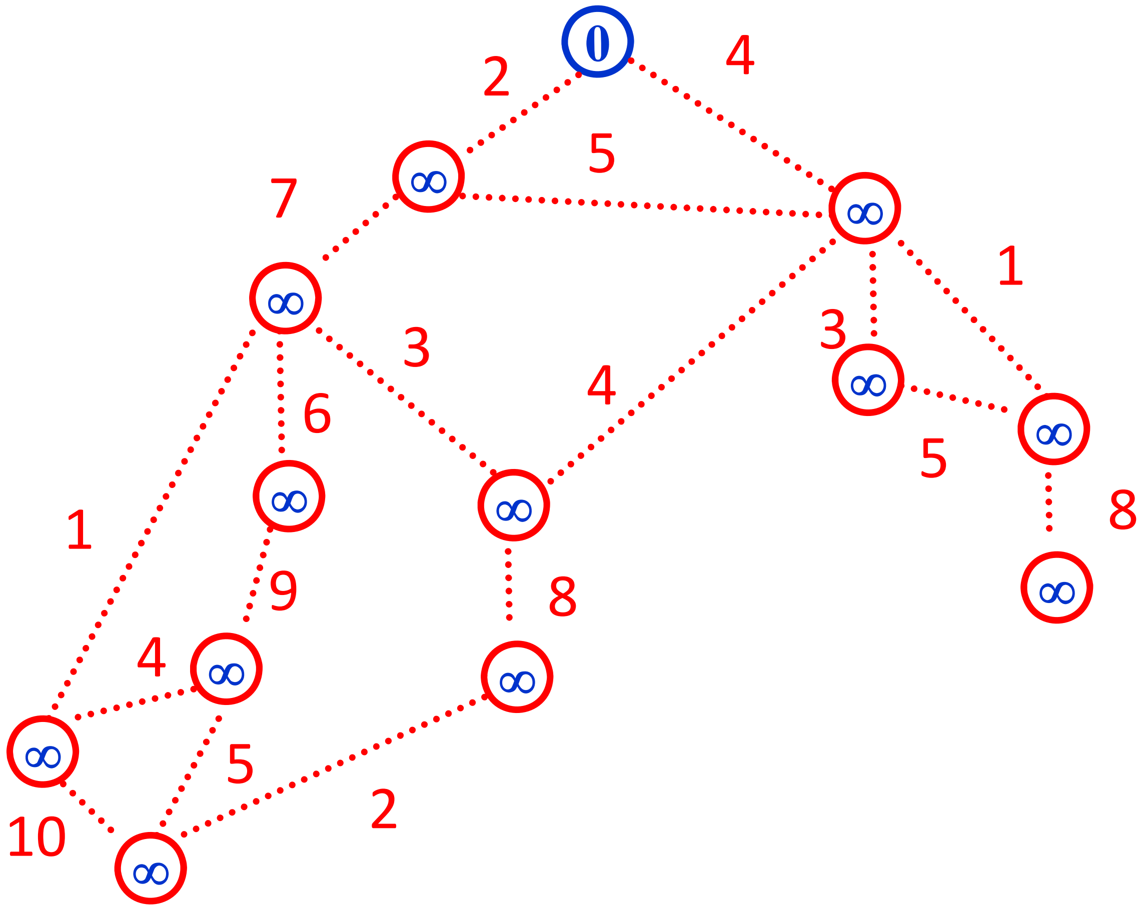
$\boldsymbol{pred}[\boldsymbol{v}] \leftarrow \boldsymbol{u}$

}

*For each $\boldsymbol{v} \notin \boldsymbol{S}$ maintain $\boldsymbol{d'}[\boldsymbol{v}]$ = minimum value of $\boldsymbol{d}[\boldsymbol{u}] + \boldsymbol{w}(\boldsymbol{e})$
over all vertices $\boldsymbol{u} \in \boldsymbol{S}$ s.t. $\boldsymbol{e} = (\boldsymbol{u}, \boldsymbol{v})$ is in $\boldsymbol{G}$

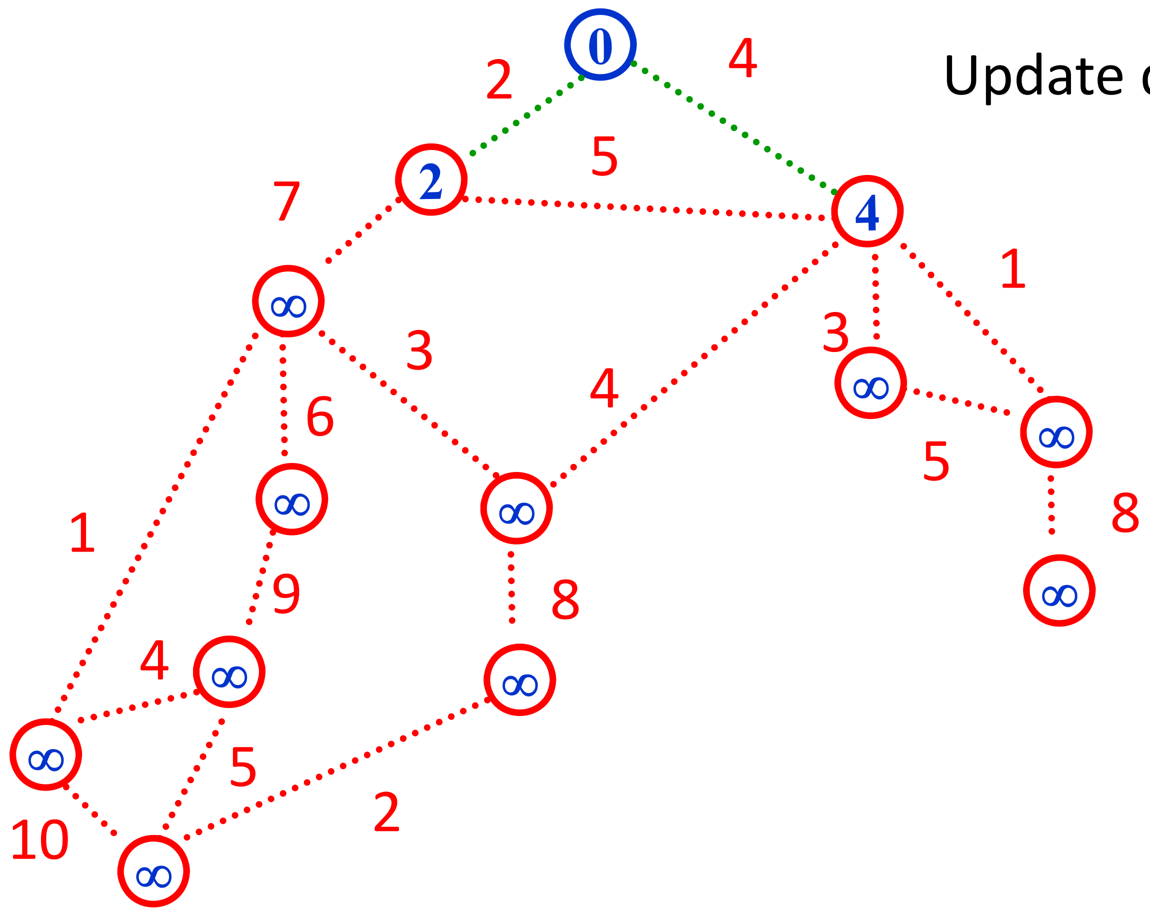
Dijkstra's Algorithm

Add to S



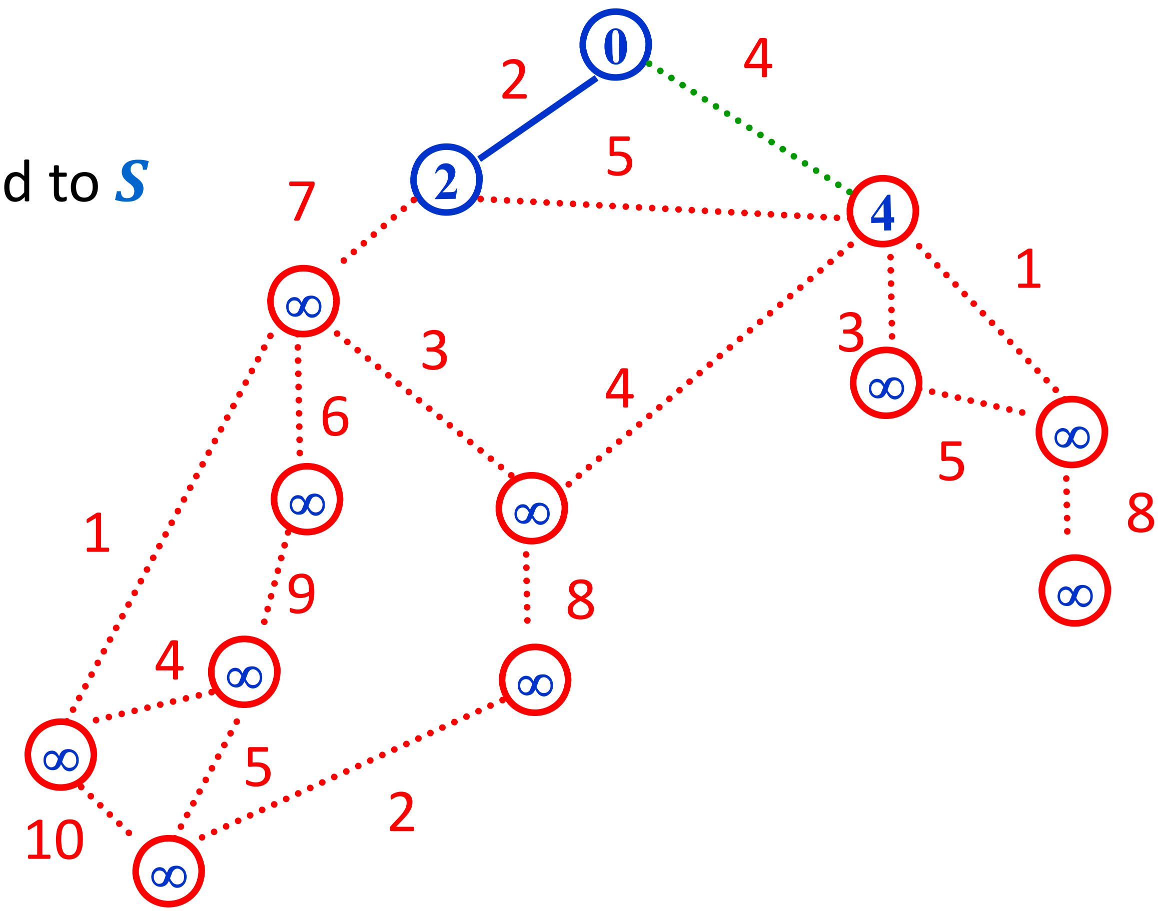
Dijkstra's Algorithm

Update distances



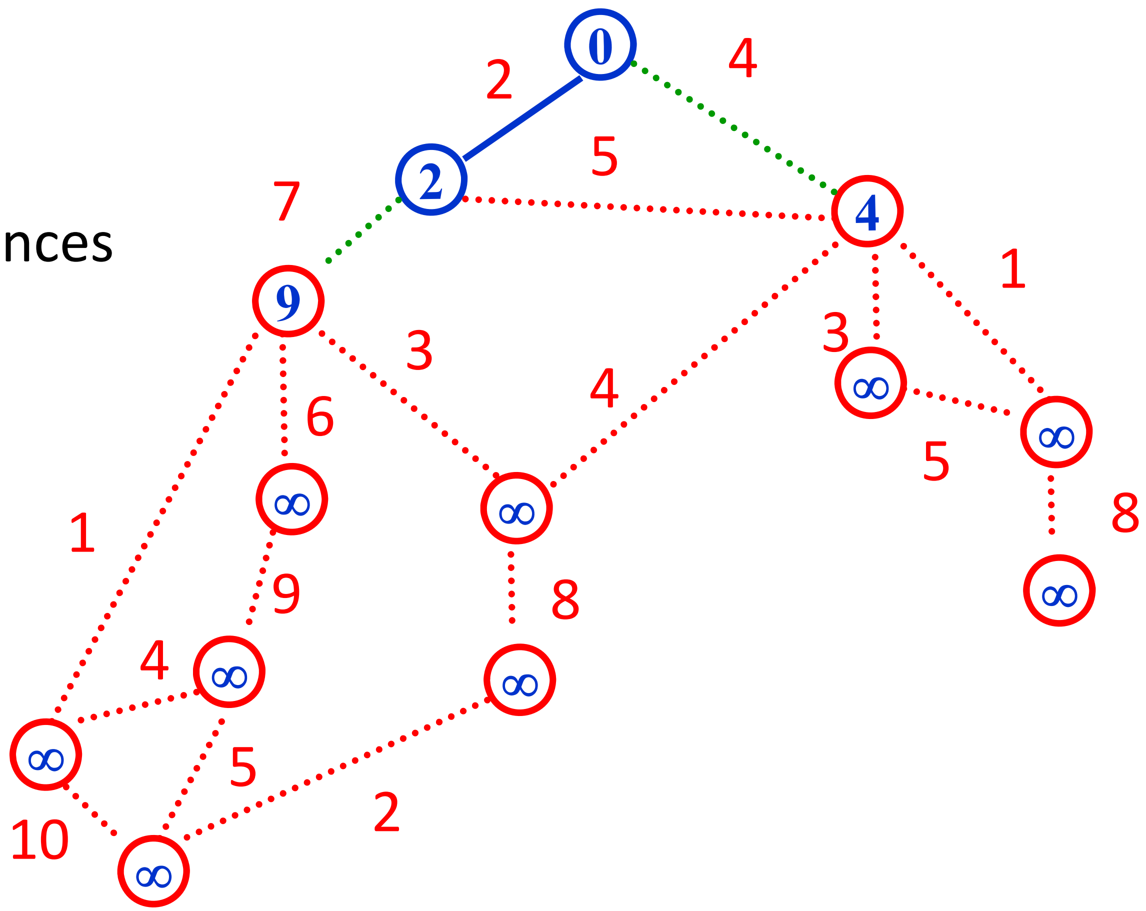
Dijkstra's Algorithm

Add to ***S***

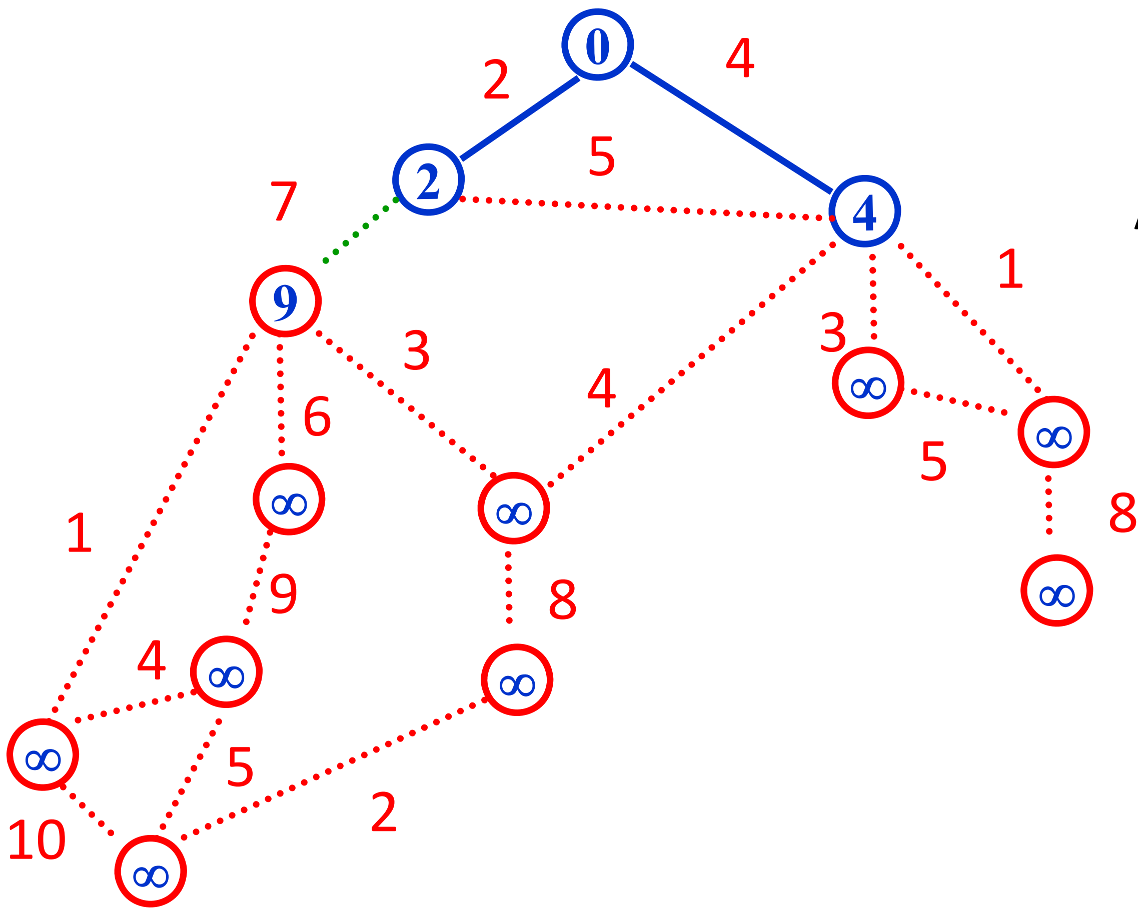


Dijkstra's Algorithm

Update distances

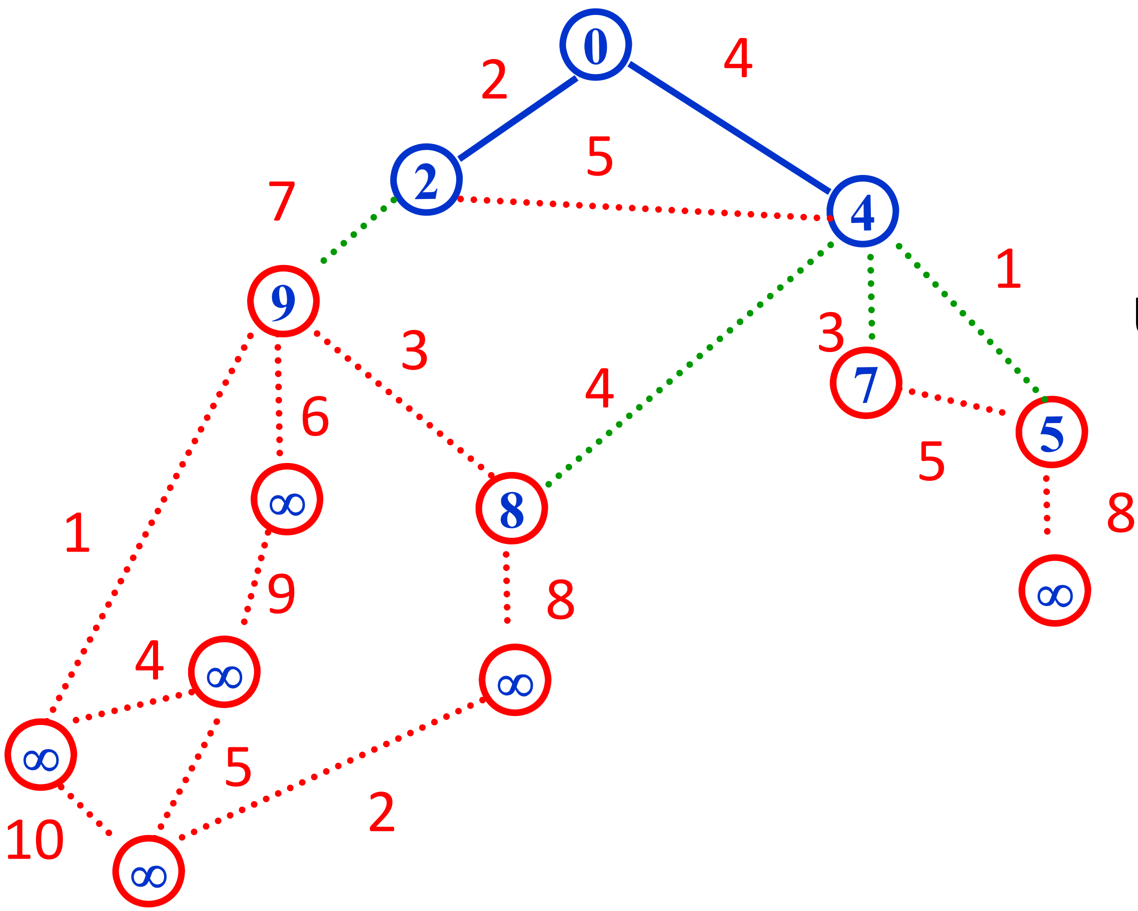


Dijkstra's Algorithm



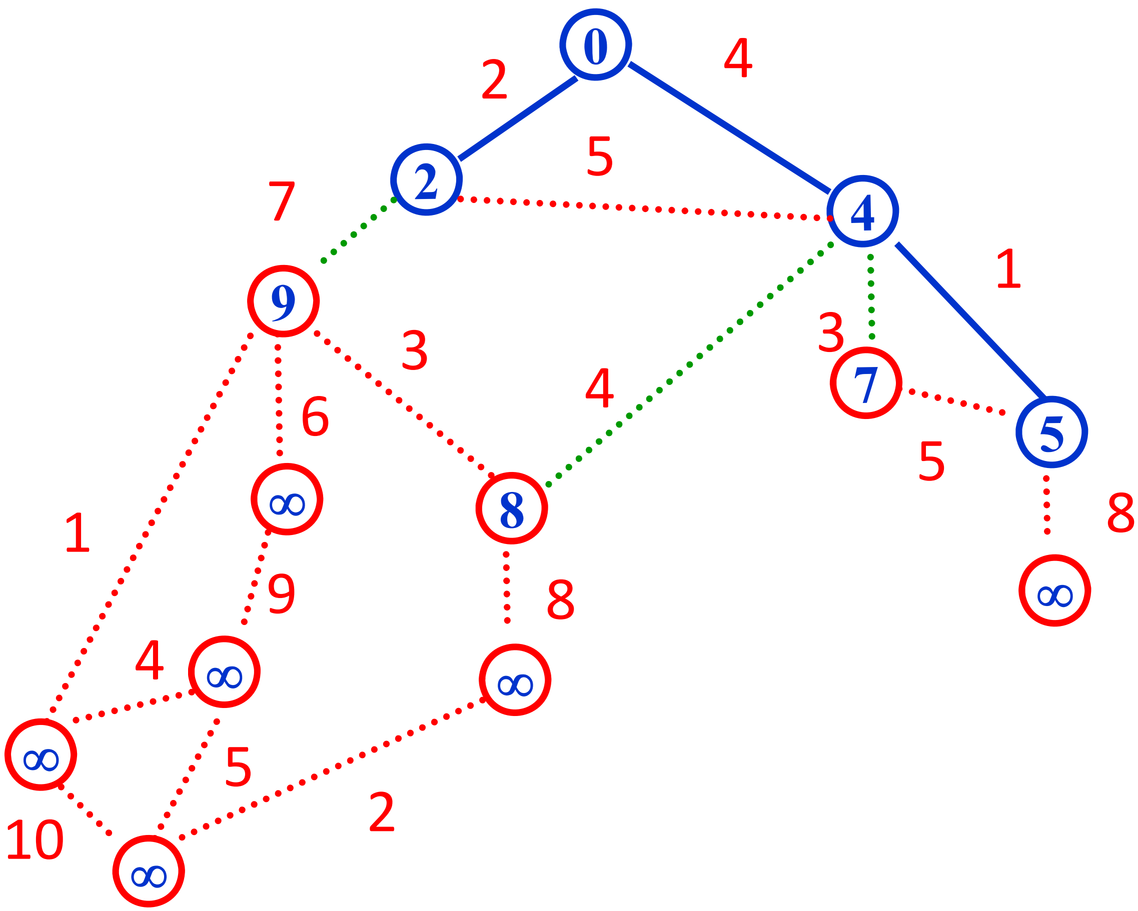
Add to *S*

Dijkstra’s Algorithm



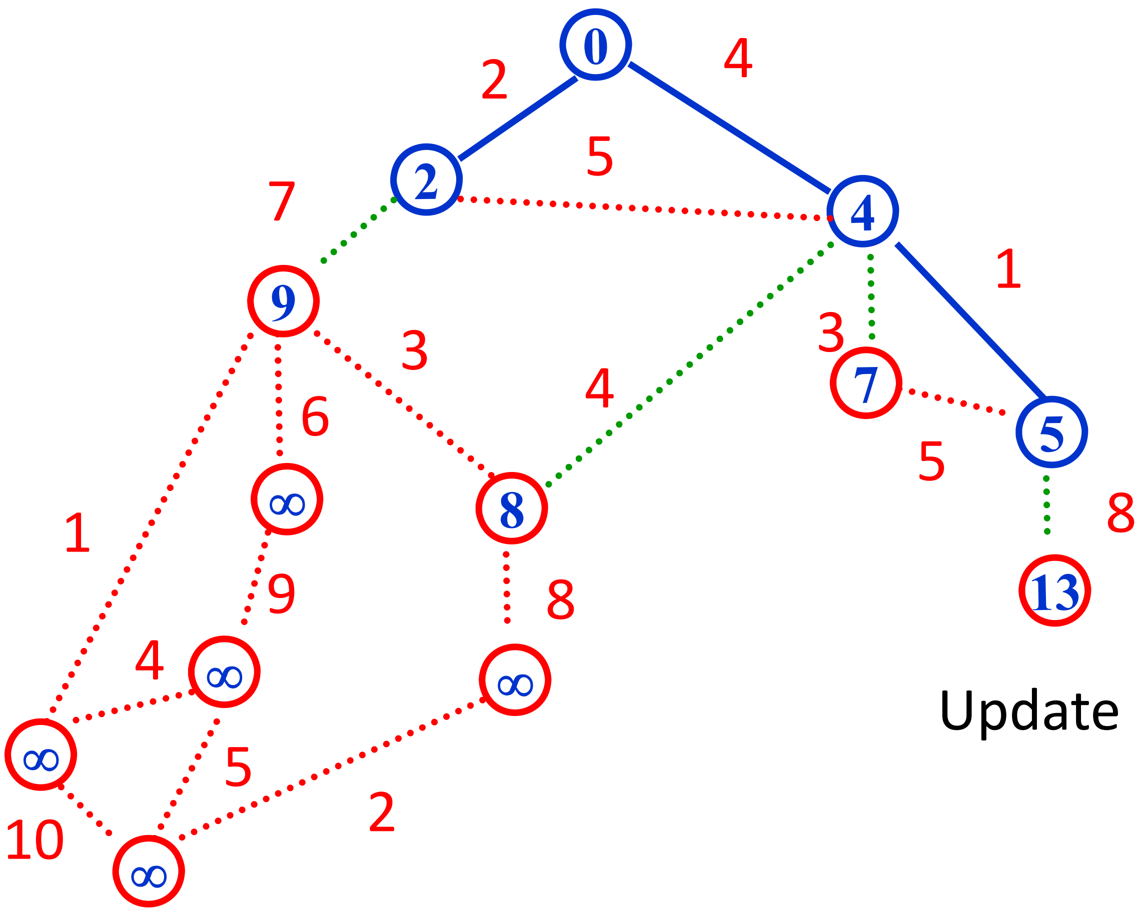
Update distances

Dijkstra's Algorithm



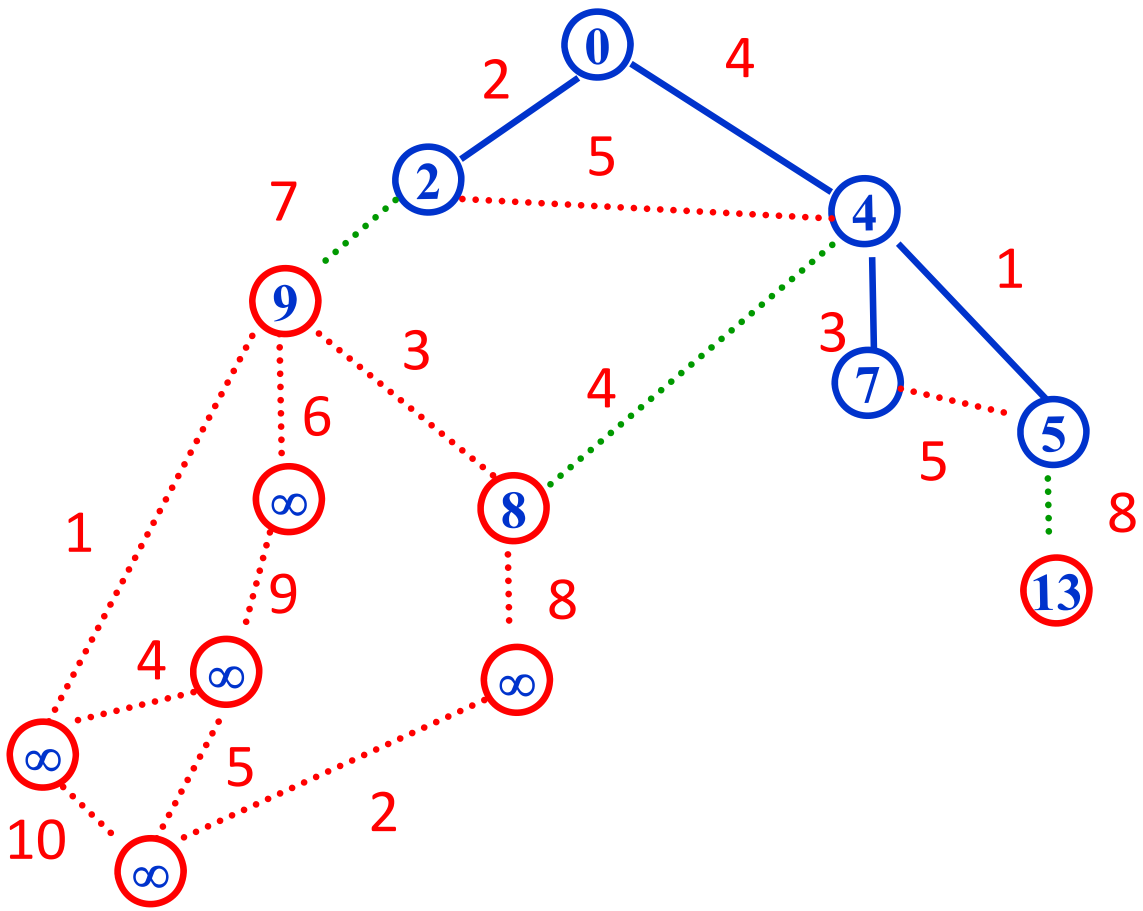
Add to S

Dijkstra's Algorithm



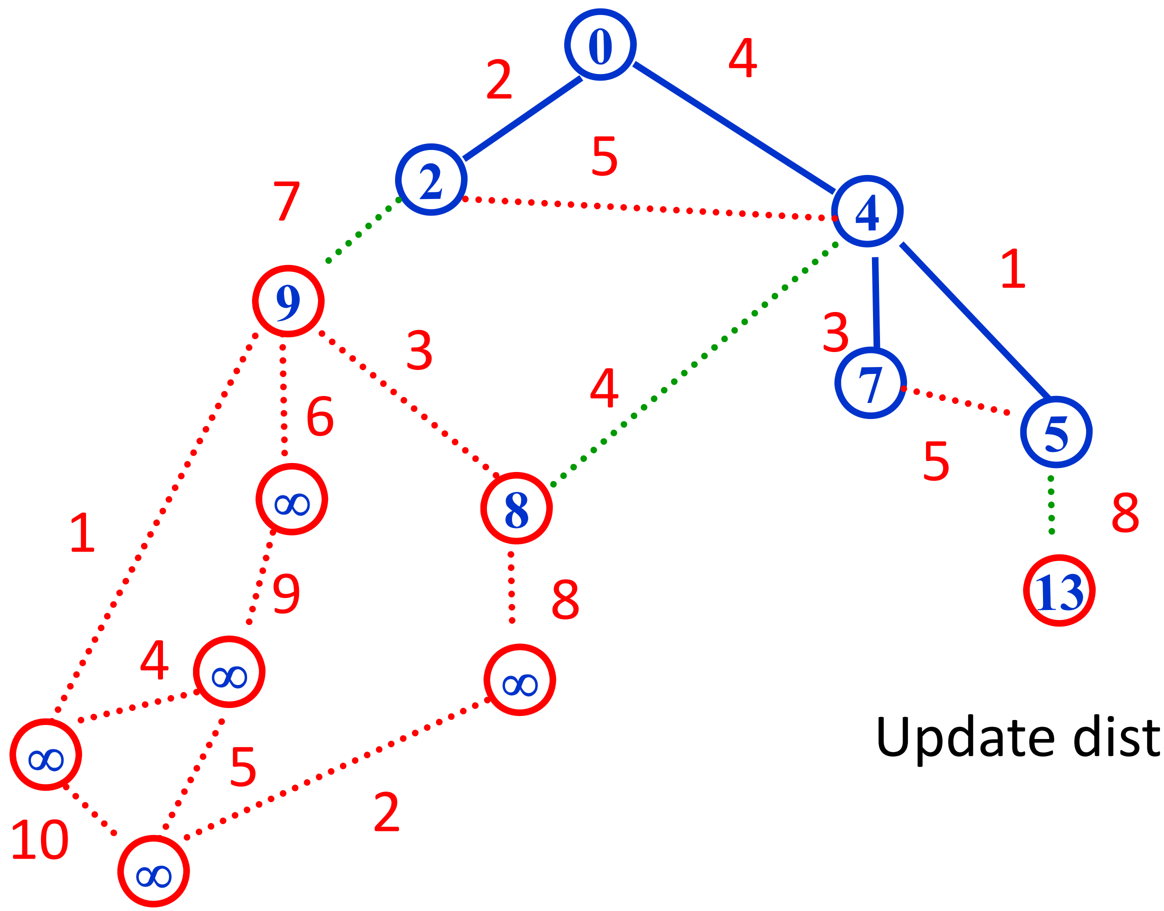
Update distances

Dijkstra's Algorithm



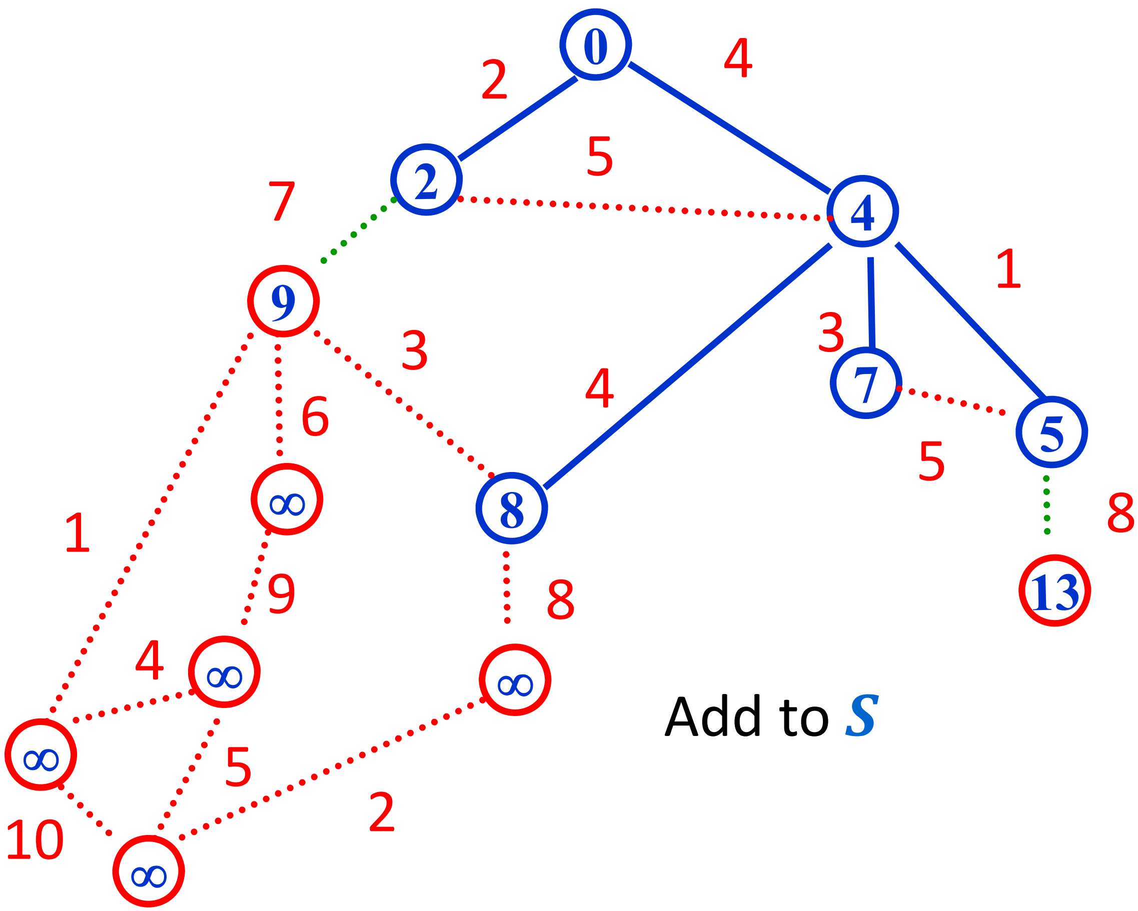
Add to **S**

Dijkstra's Algorithm



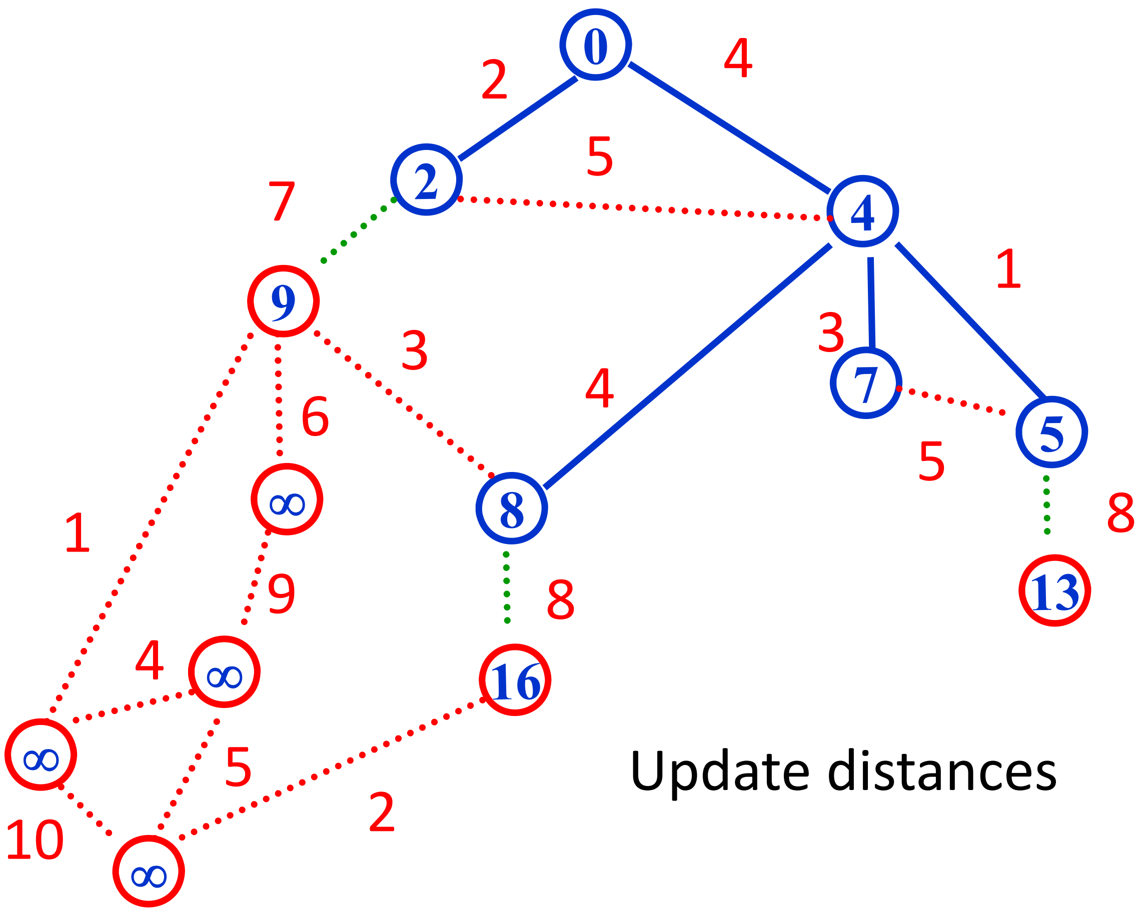
Update distances

Dijkstra's Algorithm



Add to S

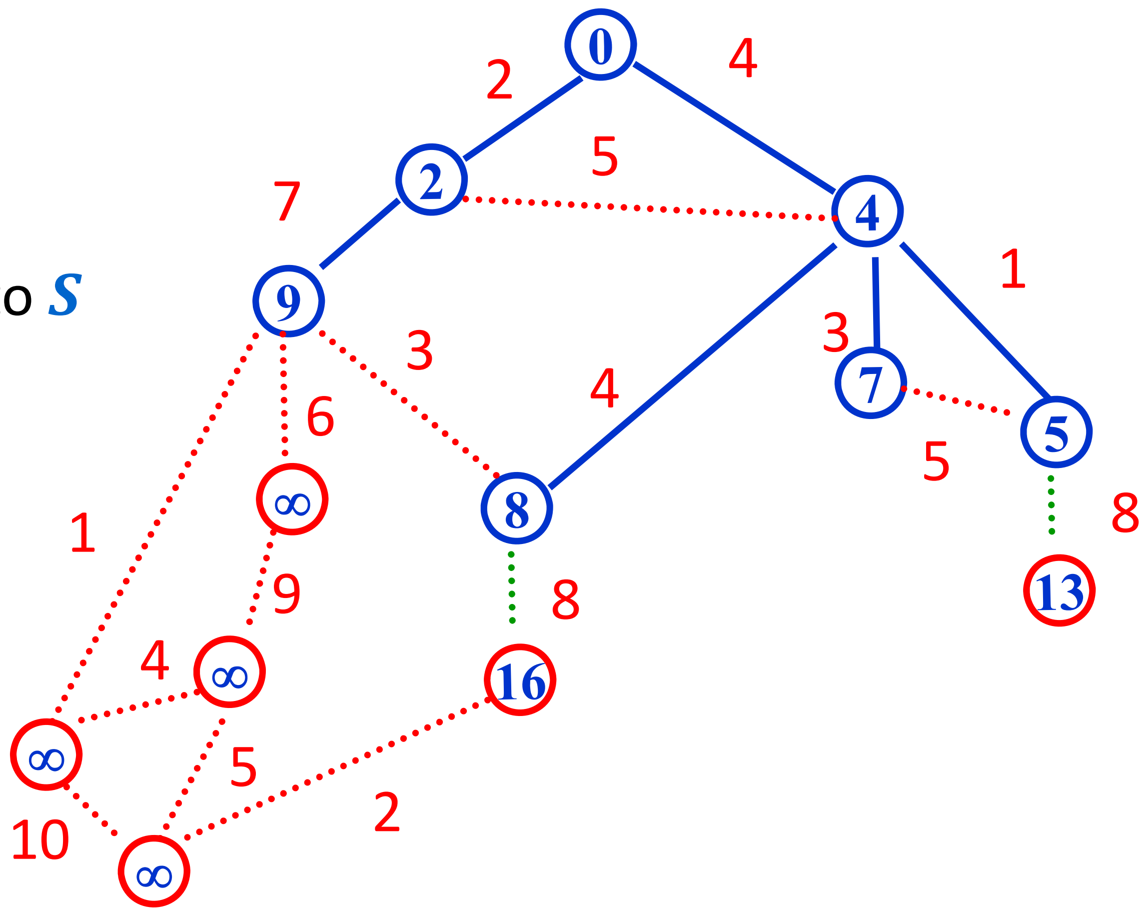
Dijkstra's Algorithm



Update distances

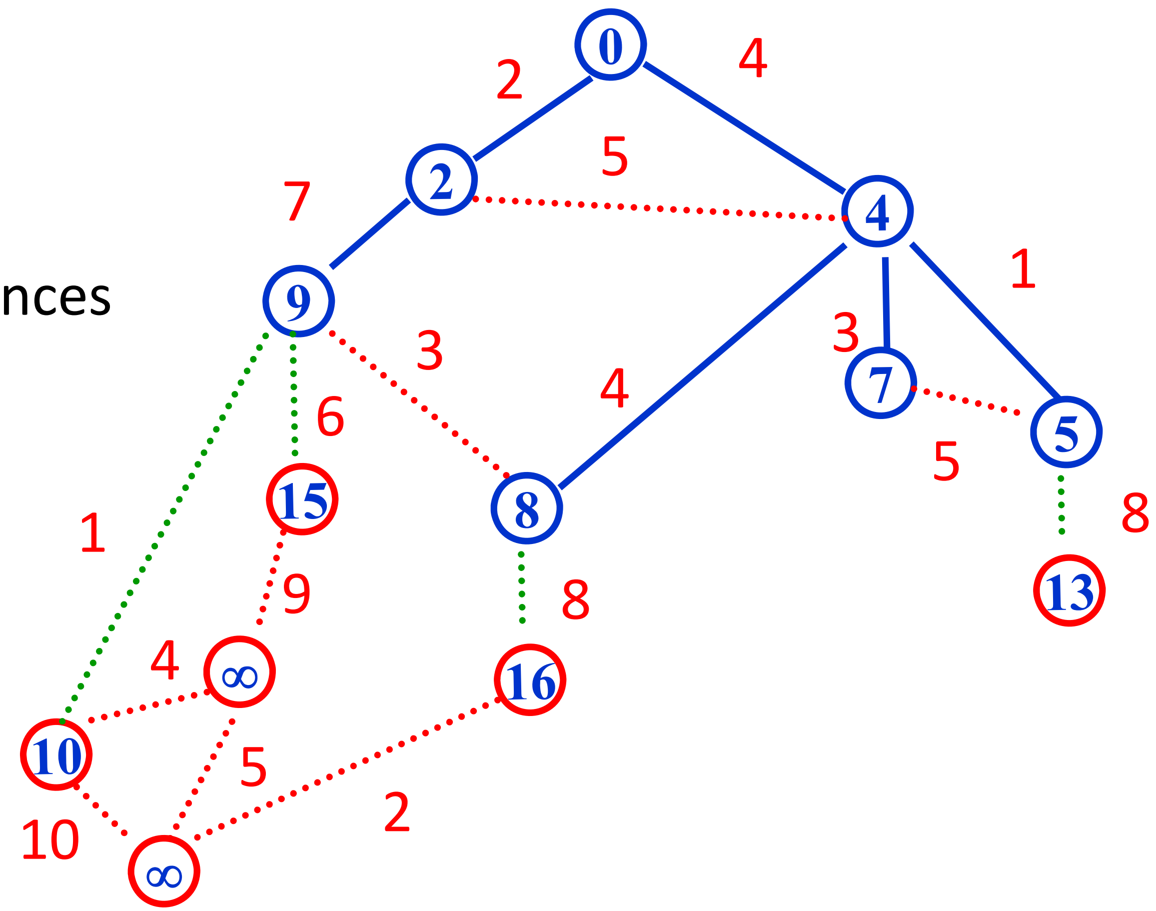
Dijkstra's Algorithm

Add to ***S***

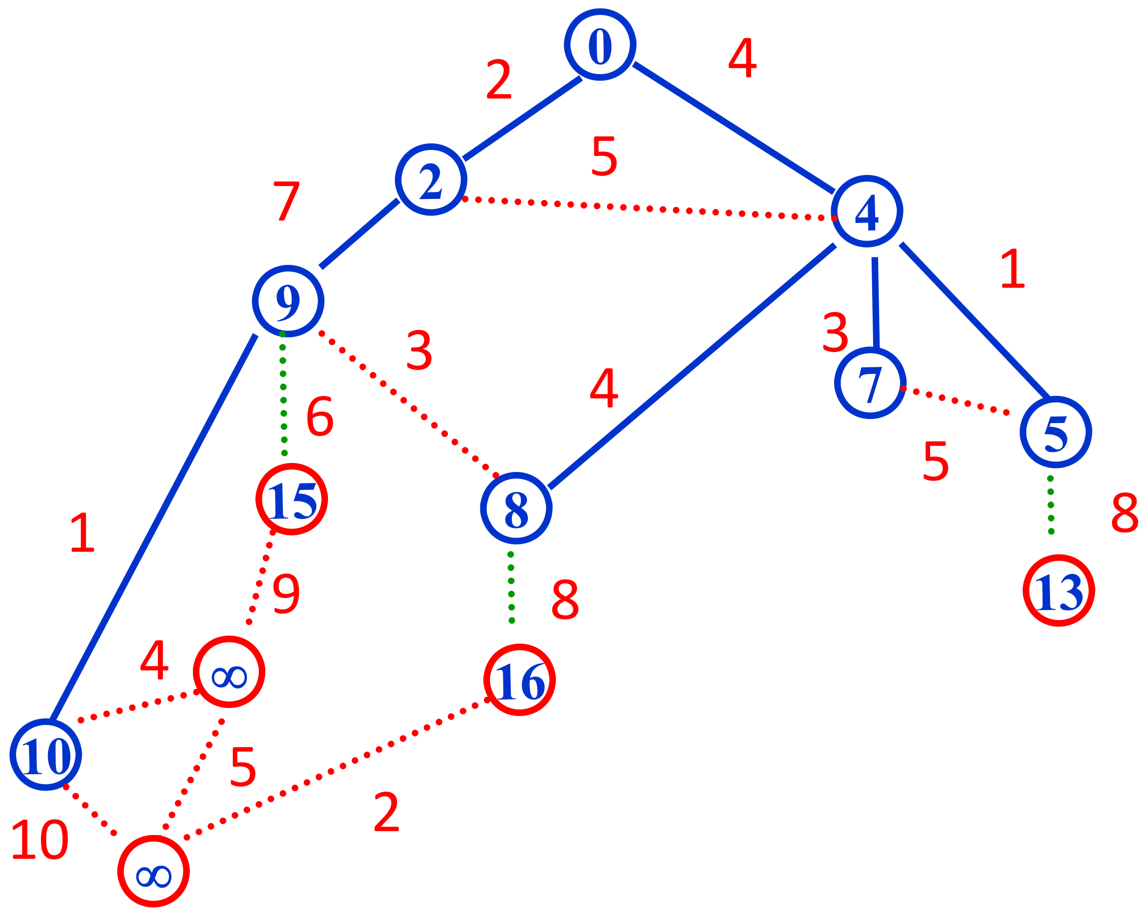


Dijkstra's Algorithm

Update distances

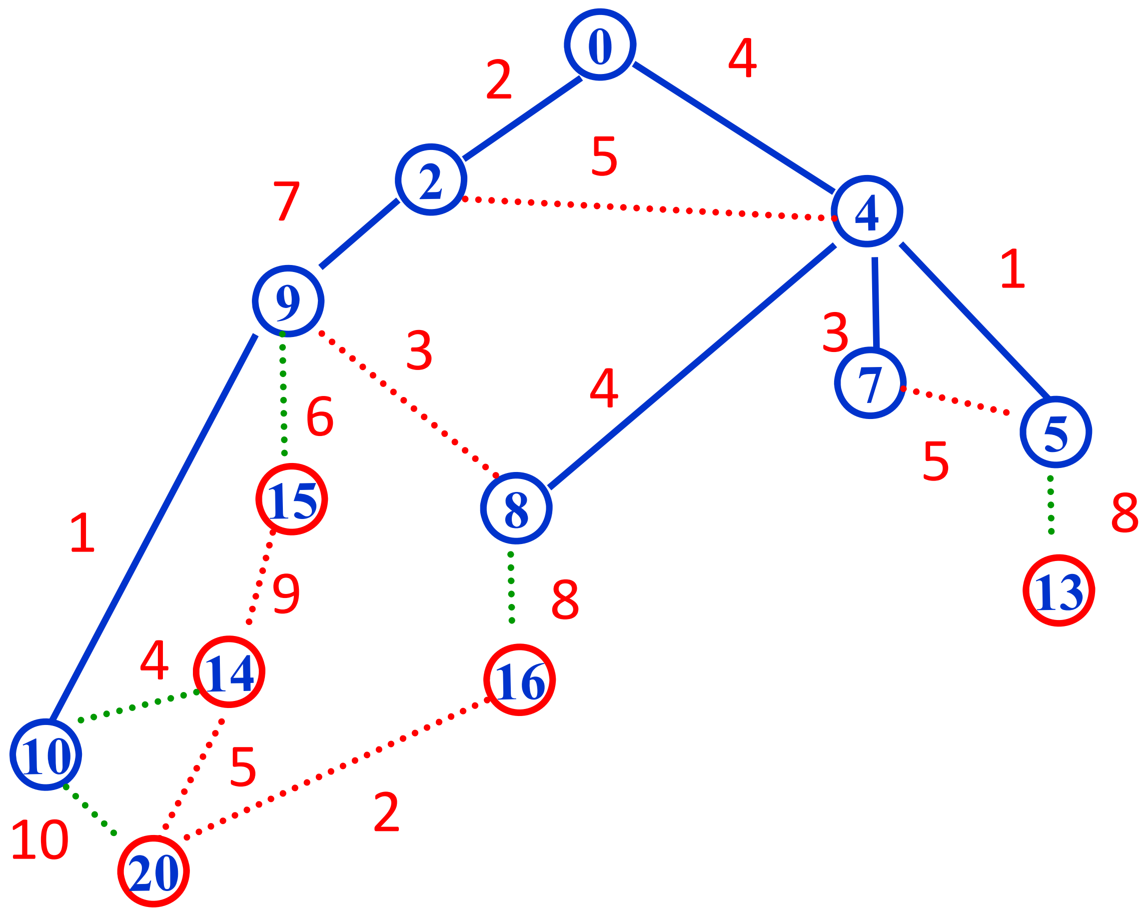


Dijkstra's Algorithm



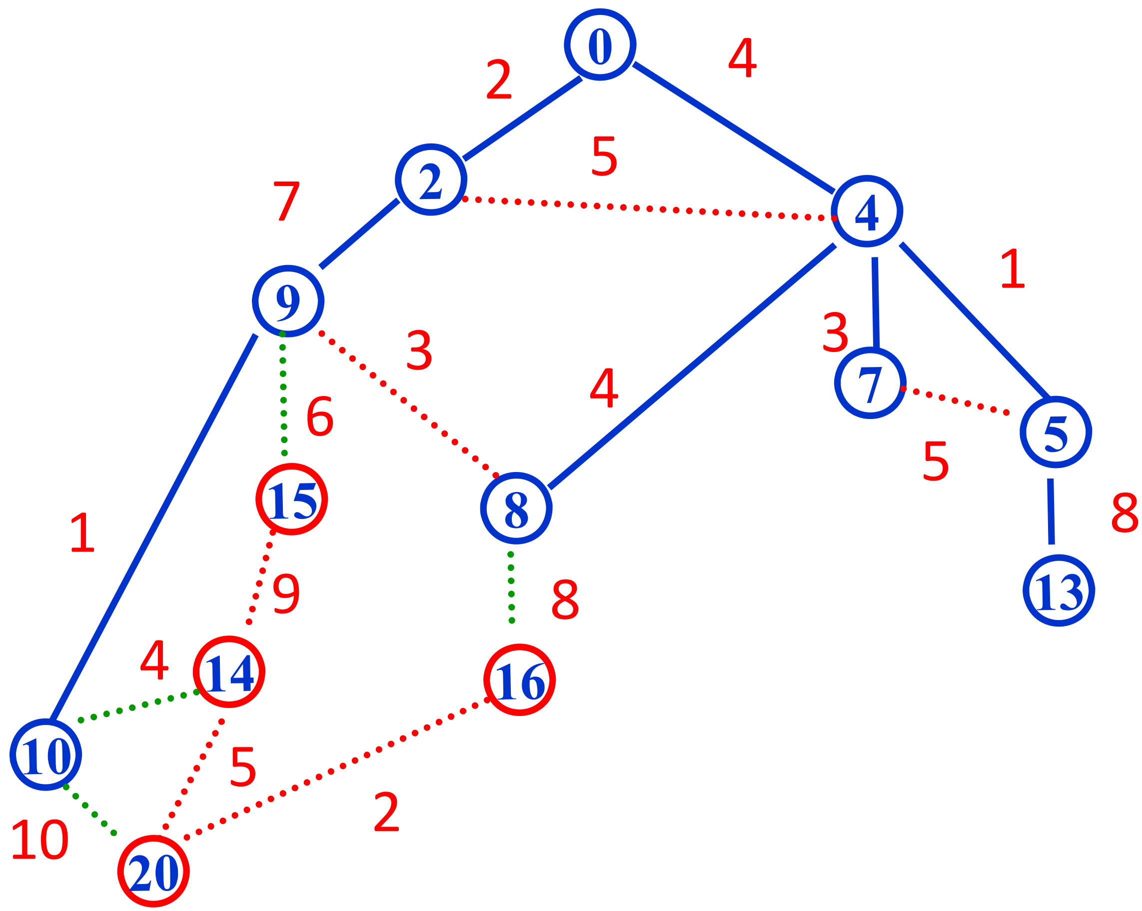
Add to ***S***

Dijkstra's Algorithm



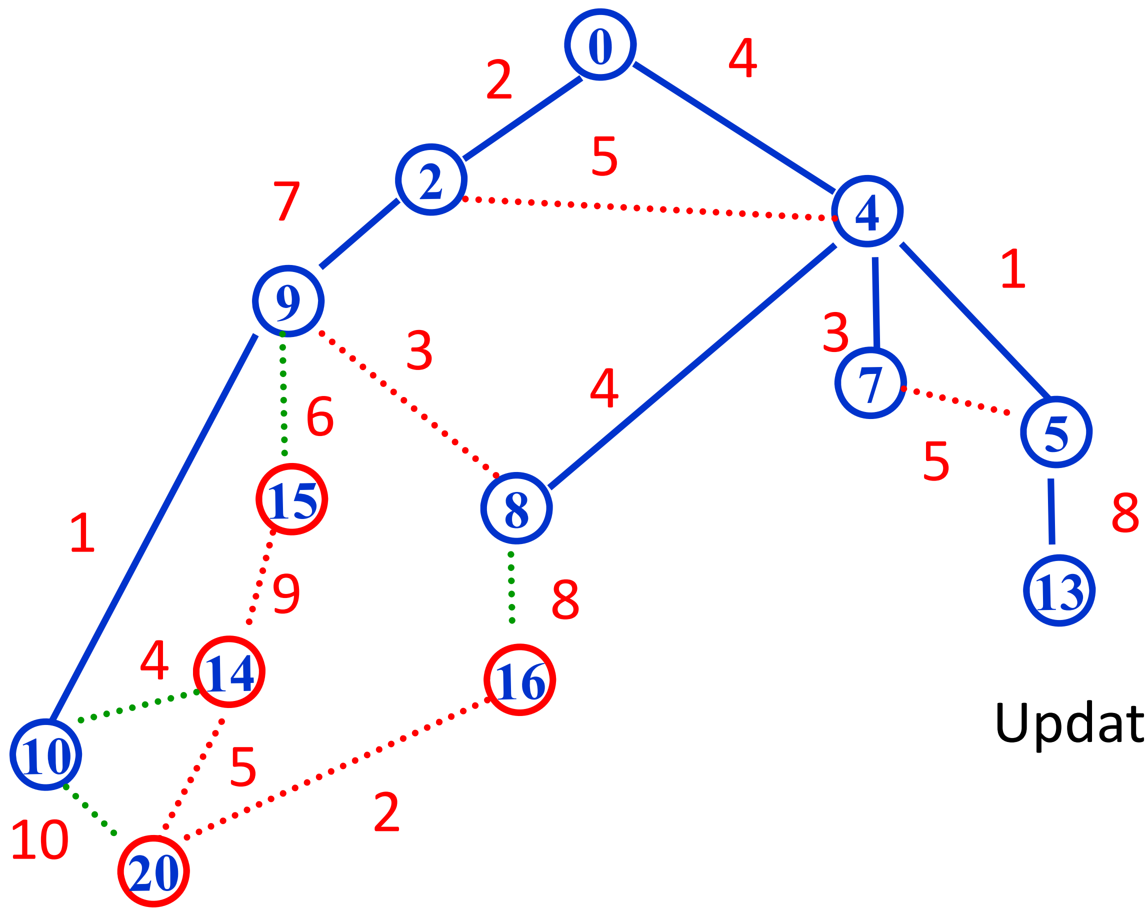
Update distances

Dijkstra's Algorithm



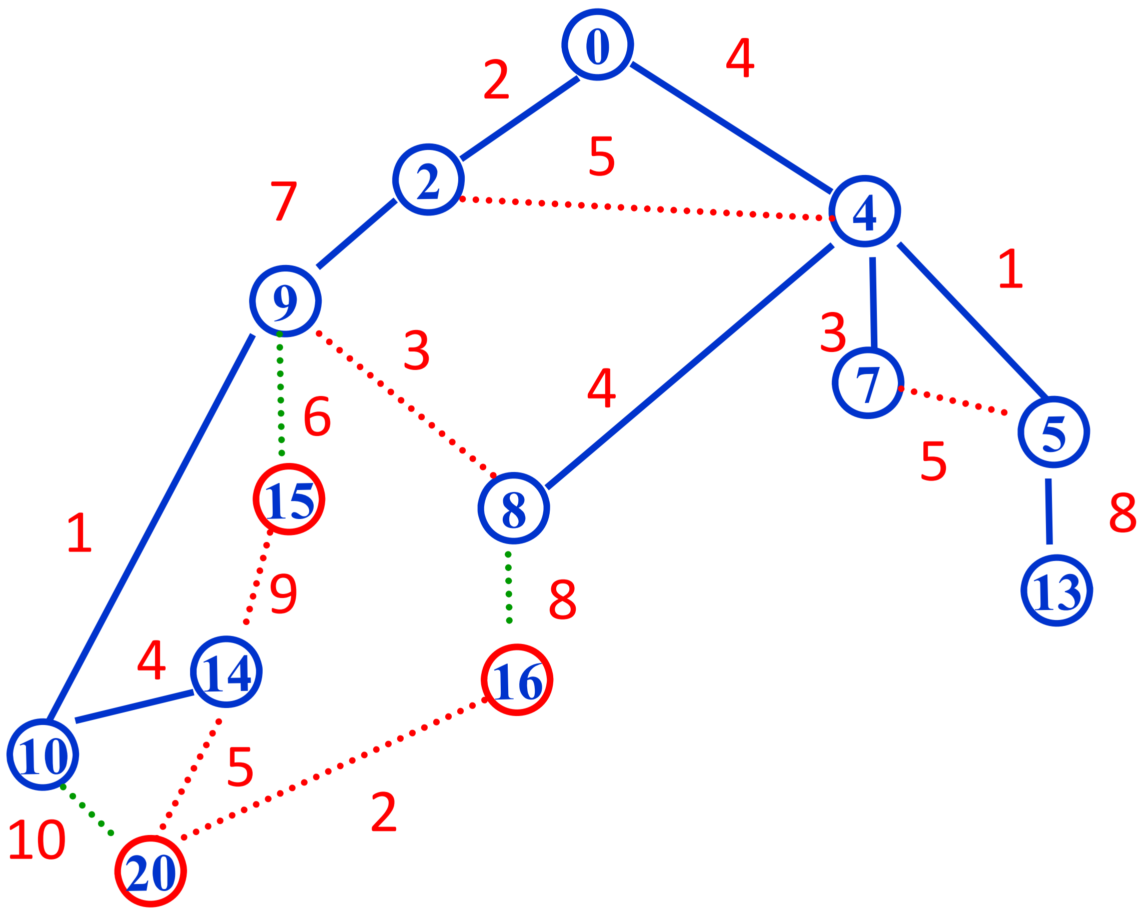
Add to *S*

Dijkstra's Algorithm



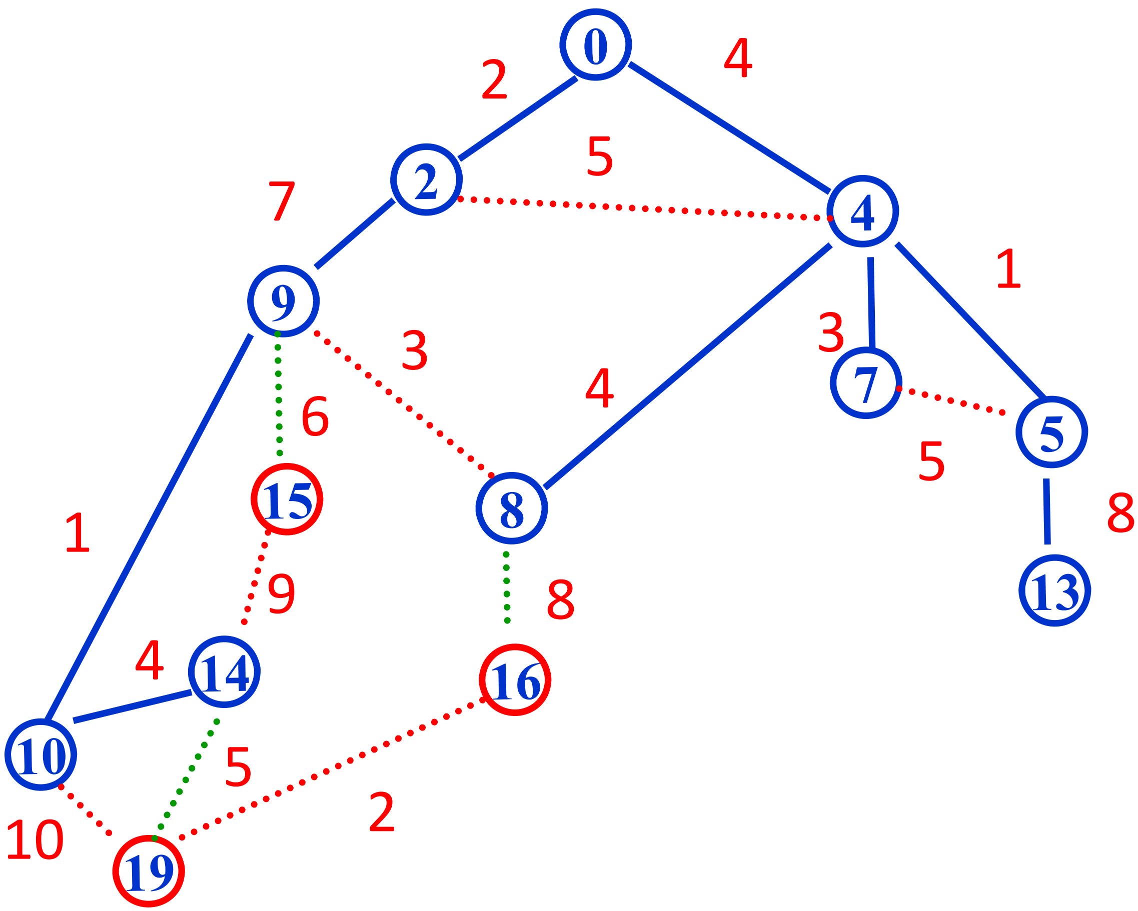
Update distances

Dijkstra's Algorithm



Add to ***S***

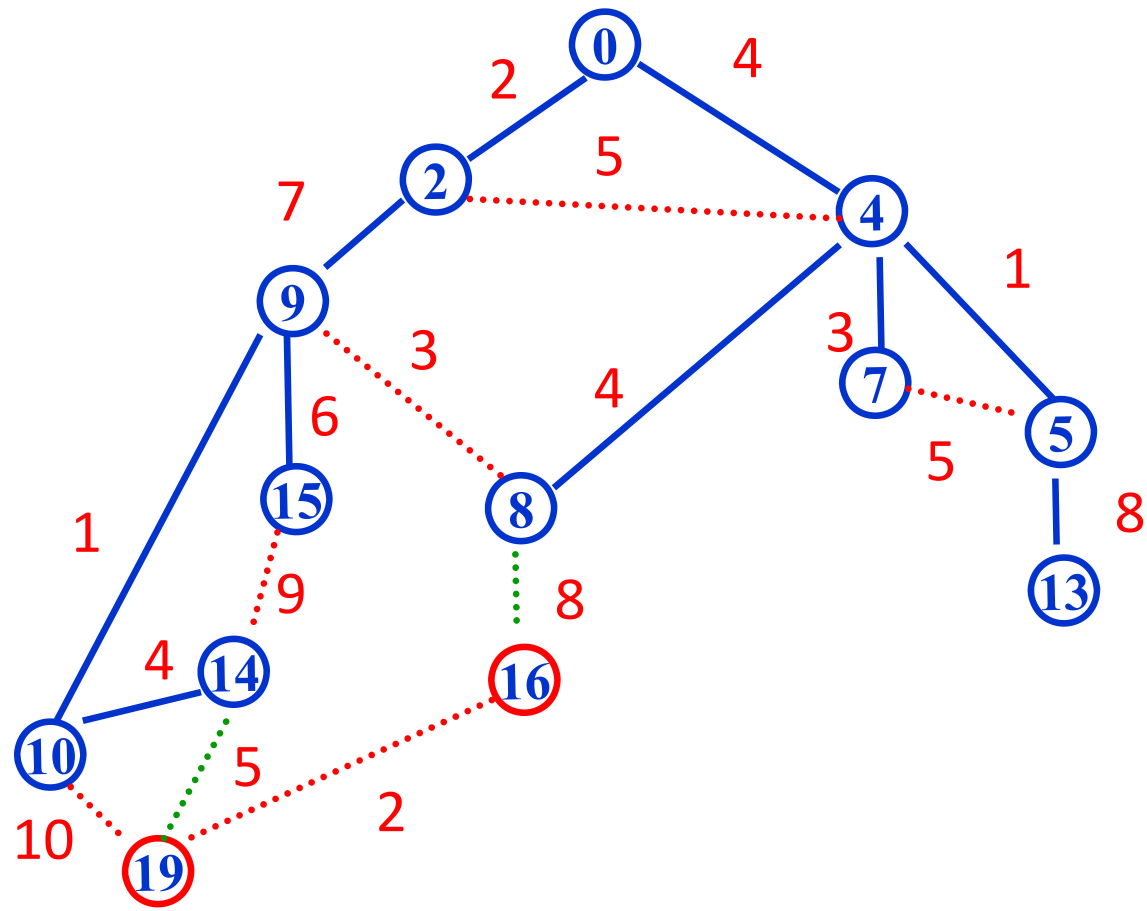
Dijkstra's Algorithm



Update distances

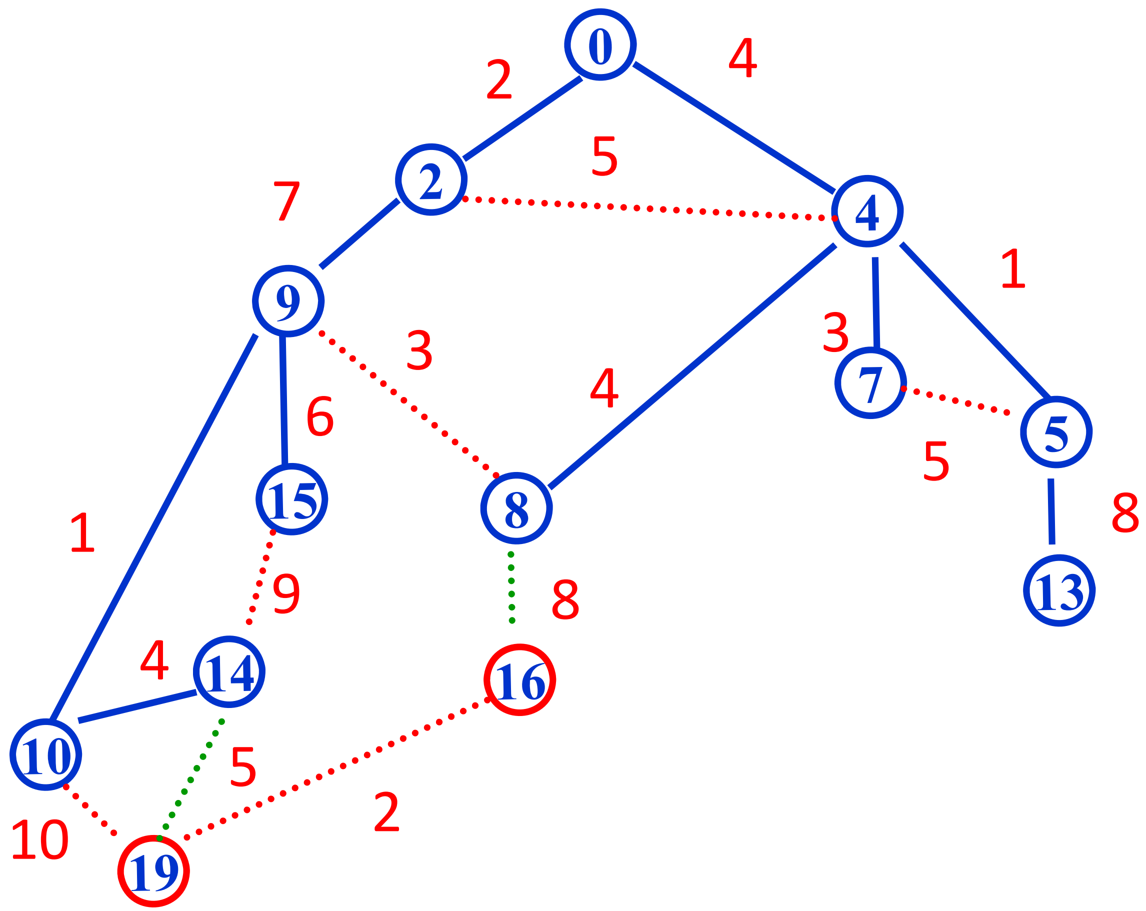
Dijkstra's Algorithm

Add to ***S***

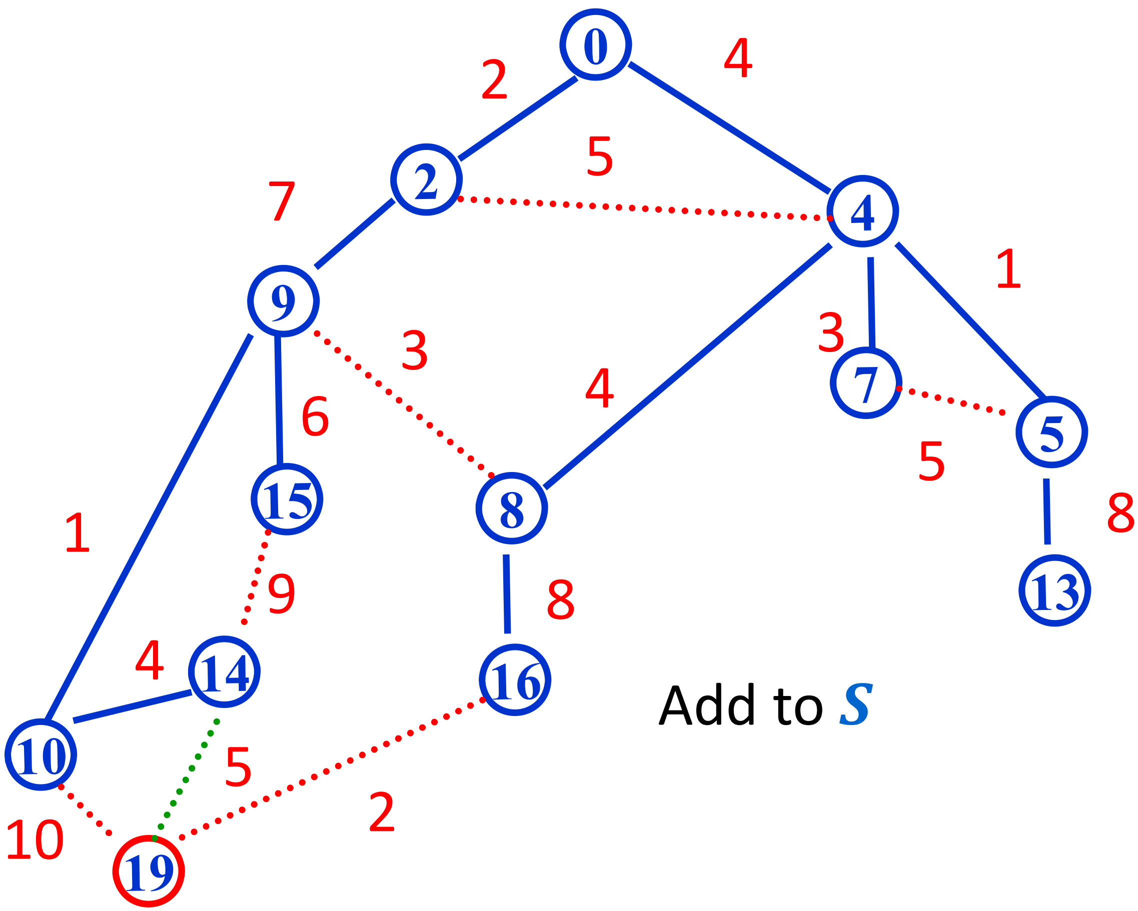


Dijkstra's Algorithm

Update distances

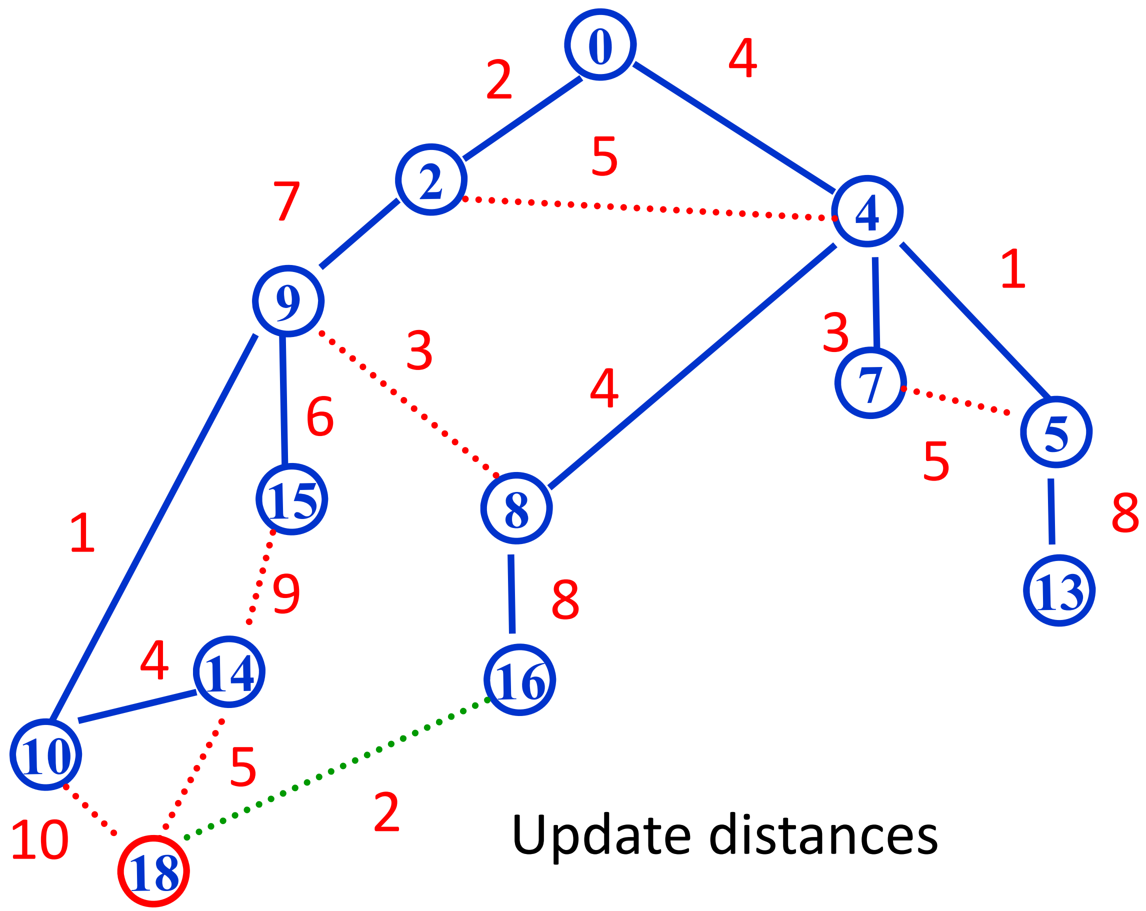


Dijkstra's Algorithm



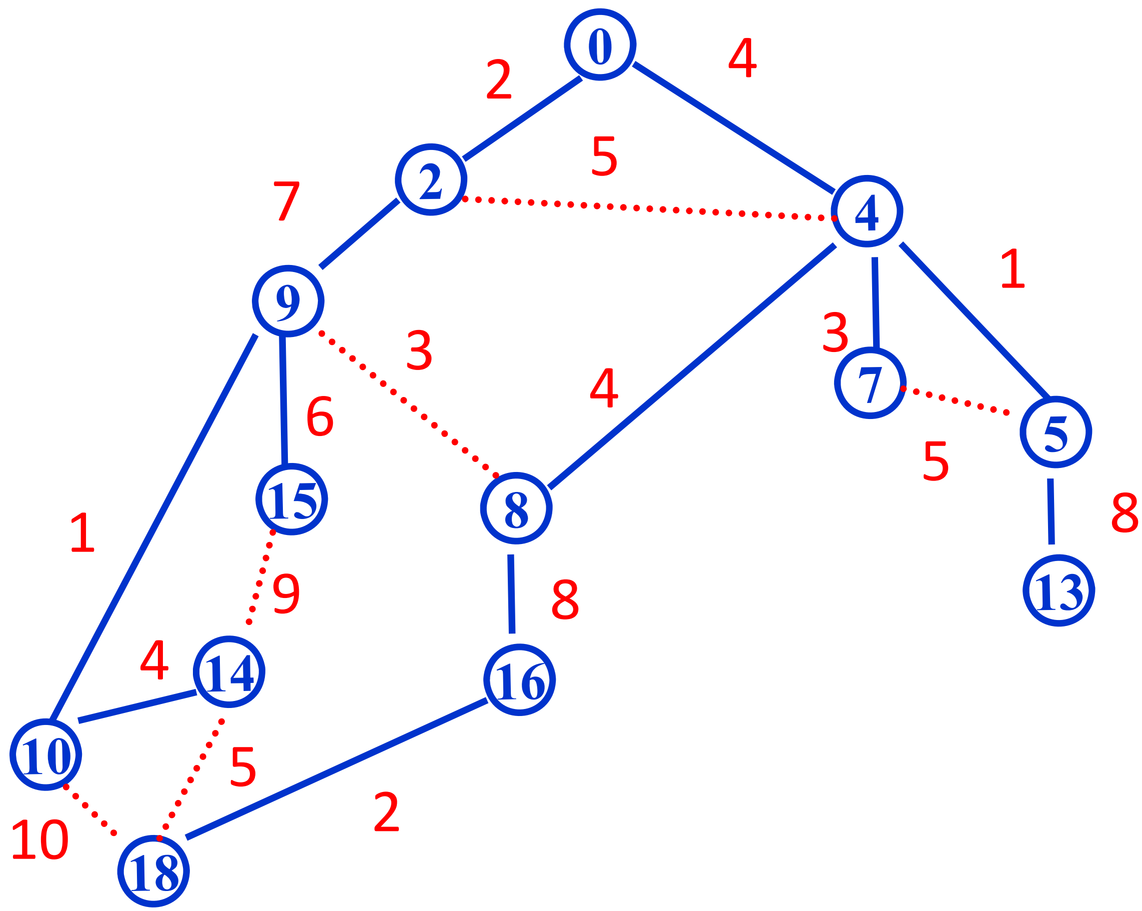
Add to *S*

Dijkstra's Algorithm



Update distances

Dijkstra’s Algorithm

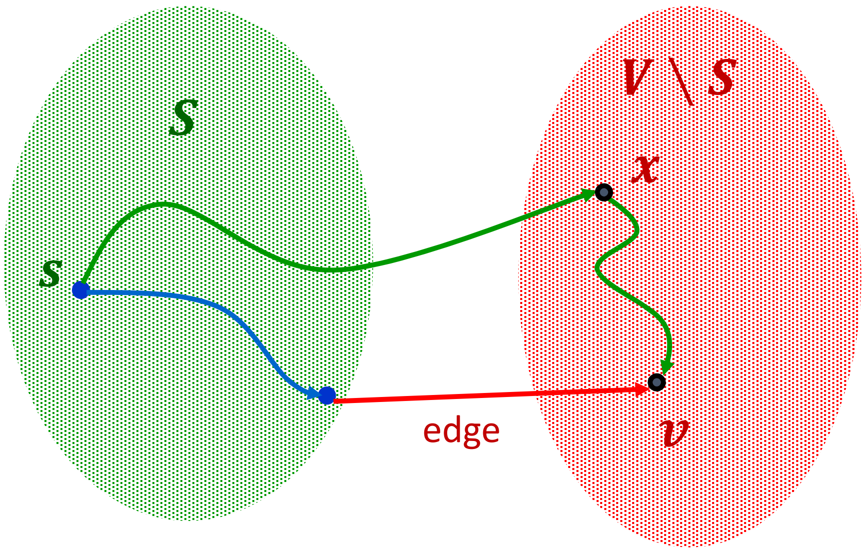


Add to ***S***

Dijkstra's Algorithm Correctness

Suppose that all distances to vertices in S are correct
and v has smallest current value $d'[v]$ in $V \setminus S$

$\Rightarrow$ $d'[v]$ = length of shortest path from s to v with only last edge leaving S



Suppose some other path P to v.
Let x = 1st vertex on this path not in S

Since v was smallest, $d'[v] \leq d'[x]$

x-v path length ≥ 0

$\Rightarrow$ length of P is at least $d'[v]$

Therefore adding v to S maintains that all distances inside S are correct

Dijkstra's Algorithm

- Algorithm also produces a tree of shortest paths to v following the inverse of $pred$ links
  - From v follow its ancestors in the tree back to s reversing edges along the path

- If all you care about is the shortest path from s to v simply stop the algorithm when v is added to S

Dijsktra’s Algorithm

Dijkstra($\boldsymbol{G}$,$\boldsymbol{w}$,s)

$\boldsymbol{S} \leftarrow \{\boldsymbol{s}\}$

$\boldsymbol{d}[\boldsymbol{s}] \leftarrow \boldsymbol{0}$

while $\boldsymbol{S} \neq \boldsymbol{V}$ {

among all edges $\boldsymbol{e} = (\boldsymbol{u}, \boldsymbol{v})$ s.t. $\boldsymbol{v} \notin \boldsymbol{S}$ and $\boldsymbol{u} \in \boldsymbol{S}$ select* one with the minimum value of $\boldsymbol{d}[\boldsymbol{u}] + \boldsymbol{w}(\boldsymbol{e})$

$\boldsymbol{S} \leftarrow \boldsymbol{S} \cup \{\boldsymbol{v}\}$

$\boldsymbol{d}[\boldsymbol{v}] \leftarrow \boldsymbol{d}[\boldsymbol{u}] + \boldsymbol{w}(\boldsymbol{e})$

$\boldsymbol{pred}[\boldsymbol{v}] \leftarrow \boldsymbol{u}$

}

*For each $\boldsymbol{v} \notin \boldsymbol{S}$ maintain $\boldsymbol{d}'[\boldsymbol{v}]$ = minimum value of $\boldsymbol{d}[\boldsymbol{u}] + \boldsymbol{w}(\boldsymbol{e})$
over all vertices $\boldsymbol{u} \in \boldsymbol{S}$ s.t. $\boldsymbol{e} = (\boldsymbol{u}, \boldsymbol{v})$ is in $\boldsymbol{G}$

Implementing Dijkstra's Algorithm

Need to

- keep current distance values $\boldsymbol{d}'[\cdot]$ for nodes in $\boldsymbol{V} \setminus \boldsymbol{S}$
- find minimum current distance value $\boldsymbol{d}'[\boldsymbol{v}]$
- reduce distances in $\boldsymbol{d}'[\cdot]$ when vertex $\boldsymbol{v}$ moved to $\boldsymbol{S}$

Data Structure Review

Priority Queue:

- Elements each with an associated **key**
- Operations
  - **Insert**
  - **Find-min**
    - Return the element with the smallest key
  - **Delete-min**
    - Return the element with the smallest key and delete it from the data structure
  - **Decrease-key**
    - Decrease the key value of some element

Implementations

- Arrays: $O(\boldsymbol{n})$ time find/delete-min, $O(\mathbf{1})$ time insert/decrease-key
- Heaps: $O(\log \boldsymbol{n})$ time insert/decrease-key/delete-min, $O(\mathbf{1})$ time find-min

Dijkstra's Algorithm with Priority Queues

- For each vertex v not in tree maintain cost $d'[v]$ of current cheapest path through tree to v
  - Store v in priority queue with key = length of this path
- Operations:
  - $n - 1$ insertions (each vertex added once)
  - $n - 1$ delete-mins (each vertex deleted once)
    - pick the vertex of smallest key, remove it from the priority queue and add its edge to the graph
  - $< m$ decrease-keys (each edge updates one vertex)

Dijskstra's Algorithm with Priority Queues

Priority queue implementations

- Array
  - insert $O(\mathbf{1})$, delete-min $O(\boldsymbol{n})$, decrease-key $O(\mathbf{1})$
  - total $O(\boldsymbol{n} + \boldsymbol{n}^{\mathbf{2}} + \boldsymbol{m}) = O(\boldsymbol{n}^2)$
- Heap
  - insert, delete-min, decrease-key all $O(\log \boldsymbol{n})$
  - total $O(\boldsymbol{m} \log \boldsymbol{n})$

Worse if $\boldsymbol{m} = \boldsymbol{\Theta}(\boldsymbol{n}^2)$

- $\boldsymbol{d}$-Heap ($\boldsymbol{d} = \boldsymbol{m}/\boldsymbol{n}$)

$\boldsymbol{m}$
  - insert, decrease-key $O(\log_{\boldsymbol{m}/\boldsymbol{n}} \boldsymbol{n})$

$\boldsymbol{n} - \mathbf{1}$
  - delete-min $O((\boldsymbol{m}/\boldsymbol{n}) \log_{\boldsymbol{m}/\boldsymbol{n}} \boldsymbol{n})$
  - total $O(\boldsymbol{m} \log_{\boldsymbol{m}/\boldsymbol{n}} \boldsymbol{n})$

Better for all values of $\boldsymbol{m}$