CS 261: Data Structures

Week 5: Priority queues

Lecture 5a: Binary and k-ary heaps

David Eppstein
University of California, Irvine

Spring Quarter, 2023

Priority queues

Example: Dijkstra's algorithm

To find the shortest path distance from a starting vertex s to each other vertex t in a directed graph with positive edge lengths:

1. Initialize dictionary D of distances; $D[s] = 0$ and, for all other vertices t, $D[t] = +\infty$
2. Initialize collection Q of not-yet-processed vertices (initially all vertices)
3. While Q is non-empty:
   - Find the vertex $v \in Q$ with the smallest value of $D[v]$
   - Remove v from Q
   - For each edge $v \to w$, set $D[w] = \min\big(D[w], D[v] + \text{length}(v \to w)\big)$

Same algorithm can find paths themselves (not just distances) by storing, for each w, its predecessor on its shortest path: the vertex v that gave the minimum value in the calculation of $D[w]$

Priority queue operations used by Dijkstra, I

For Dijkstra to be efficient, we need to organize Q into a priority queue data structure allowing vertex with minimum D to be found quickly.

First operation used by Dijkstra: Create a new priority queue

1. Initialize dictionary D of distances; $D[s] = 0$ and, for all other vertices t, $D[t] = +\infty$
2. Initialize collection Q of not-yet-processed vertices (initially all vertices)
3. While Q is non-empty:
   - Find the vertex $v \in Q$ with the smallest value of $D[v]$
   - Remove v from Q
   - For each edge $v \to w$, set $D[w] = \min\big(D[w], D[v] + \text{length}(v \to w)\big)$

Priority queue operations used by Dijkstra, II

Second operation used by Dijkstra:
Find and remove item with minimum value

(Some applications use maximum value; some descriptions of this data structure separate find and remove into two operations.)

1. Initialize dictionary D of distances; $D[s] = 0$ and, for all other vertices t, $D[t] = +\infty$
2. Initialize collection Q of not-yet-processed vertices (initially all vertices)
3. While Q is non-empty:
   - Find the vertex $v \in Q$ with the smallest value of $D[v]$
   - Remove v from Q
   - For each edge $v \to w$, set $D[w] = \min\big(D[w], D[v] + \text{length}(v \to w)\big)$

Priority queue operations used by Dijkstra, III

Third operation used by Dijkstra: Change item's priority

In Dijkstra the priority always gets smaller (earlier in the priority ordering); this will be important for efficiency.

1. Initialize dictionary D of distances; $D[s] = 0$ and, for all other vertices t, $D[t] = +\infty$
2. Initialize collection Q of not-yet-processed vertices (initially all vertices)
3. While Q is non-empty:
   - Find the vertex $v \in Q$ with the smallest value of $D[v]$
   - Remove v from Q
   - For each edge $v \to w$, set $D[w] = \min\big(D[w], D[v] + \text{length}(v \to w)\big)$

Summary of priority queue operations

Used by Dijkstra:

- Create new priority queue for elements with priorities
- Find and remove minimum-priority element
- Decrease the priority of an element

Also sometimes useful:

- Add an element
- Remove arbitrary element
- Increase priority
- Merge two queues

Priority queues in Python

Main built-in type: heapq

Implements binary heap (this lecture), represented as dynamic array

Operations:

- `heapify`: Create new priority queue for collection of elements (no separate priorities: ordered by standard comparisons)
- `heappop`: Find and remove minimum priority element
- `heappush`: Add an element

No ability to associate numeric priorities to arbitrary vertex objects.
No ability to change priorities of elements already in queue.

Dijkstra in Python

Cannot use decrease-priority; instead, store bigger priority queue of triples (d, u, v) where $u \to v$ is an edge and d is the length of the shortest path to u plus the length of the edge

First triple (d, u, v) found for v is the one giving the shortest path

```
def dijkstra(G,s,length):
    D = {}
    Q = [(0,None,s)]
    while Q:
        dist,u,v = heapq.heappop(Q)
        if v not in D:
            D[v] = dist
            for w in G[v]:
                heapq.heappush(Q,(D[v]+length(v,w),v,w))
    return D
```

Binary heaps

Binary heaps

Binary heaps are a type of priority queue

(Min-heap: smaller priorities are earlier; max-heap: opposite)

Based on heap-ordered trees (min is at root) and an efficient representation of binary trees by arrays

(No need for separate tree nodes and pointers

Allows priorities to change

(But to do this you need a separate data structure to keep track of the location of each object in the array – this extra complication is why Python doesn't implement this operation)

Find-and-remove, add, change-priority all take $O(\log n)$ time
Creating a binary heap from n objects takes $O(n)$ time

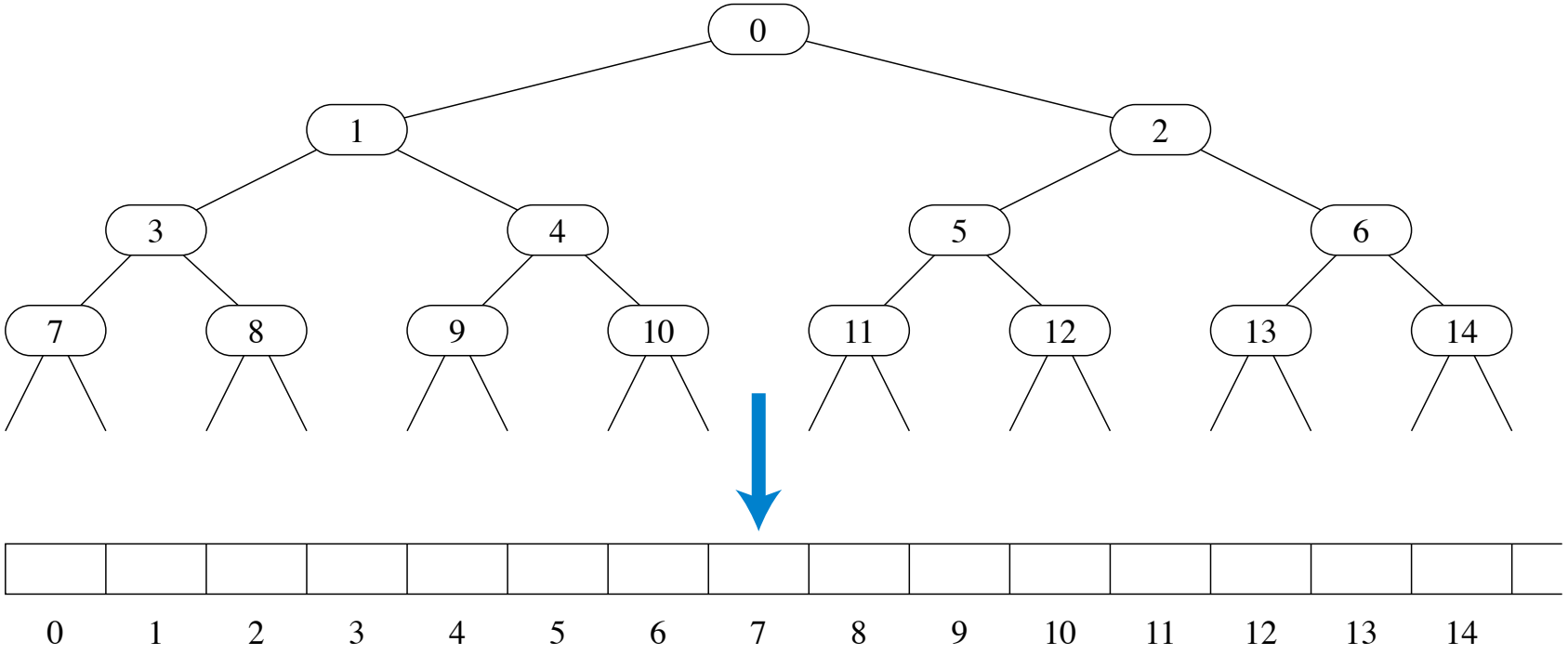
This is what Python `heapq` uses

Representing trees using arrays

Key idea: number the nodes of infinite complete binary tree row-by-row, and left-to-right within each row

Use node number as index for array cell representing that node

For finite trees with n nodes, use the first n indexes



$\text{parent}(x) = \lfloor (x-1)/2 \rfloor$ $\qquad$ $\text{children}(x) = (2x+1, 2x+2)$

Heap order

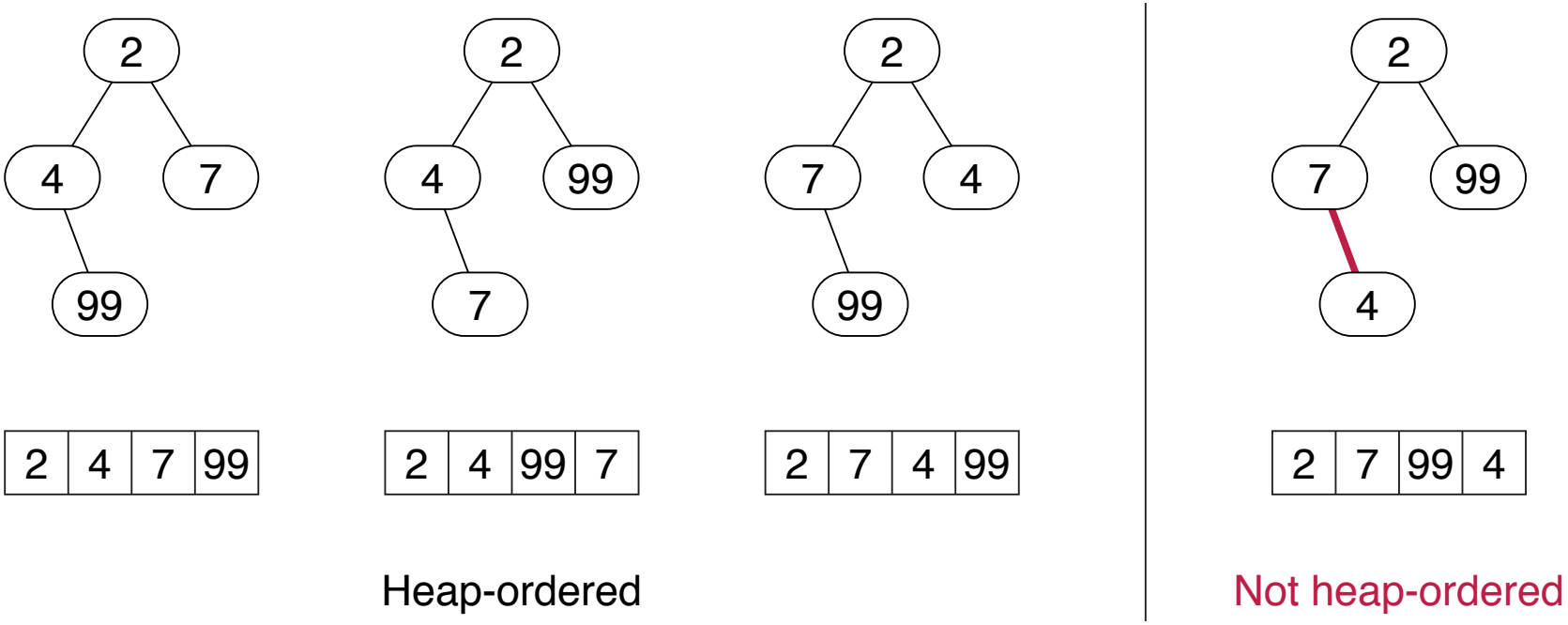
A tree is heap-ordered when, for every non-root node x,

$$\text{priority}\big(\text{parent}(x)\big) \leq \text{priority}(x)$$

A binary heap on n items is an array H of the items, so that the corresponding binary tree is heap-ordered: For all $0 < i < n$,

$$\text{priority}\left(H\left[\left\lfloor \frac{i-1}{2} \right\rfloor\right]\right) \leq \text{priority}\big(H[i]\big)$$

Examples of heap order



Sorted arrays are heap-ordered (left example)
but heap-ordered arrays might not be sorted (middle two examples)

Find and remove minimum-priority element

Where to find it: always in array cell 0

How to remove it:

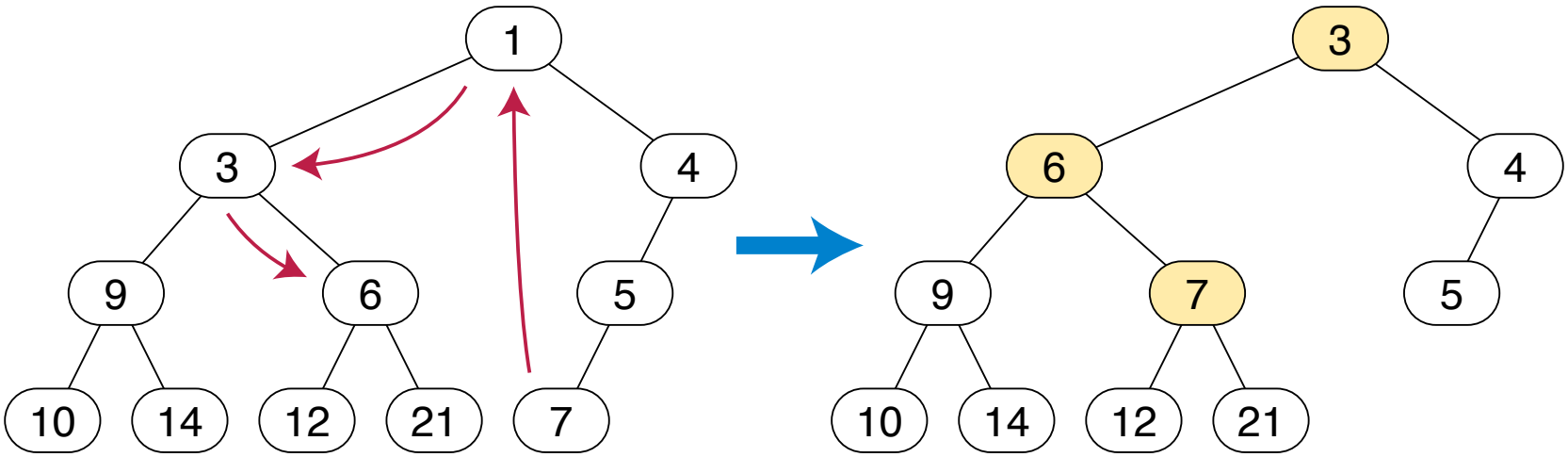
1. Swap element zero with element $n - 1$
2. Reduce array length by one
3. Fix heap at position 0 (might now be larger than children)

To fix heap at position p, repeat:

1. If p has no children ($2p + 1 \geq n$), return
2. Find child q with minimum priority
3. If priority(p) $\leq$ priority(q), return
4. Swap elements at positions p and q
5. Set $p = q$ and continue repeating

Time = $O(1)$ per level of the tree = $O(\log n)$ total

Example of delete-min operation



1	3	4	9	6	5	10	14	12	21	7

3	6	4	9	7	5	10	14	12	21

To delete the minimum element (element 1):

Swap the final element (7) into root position and shrink array

Swap element 7 with its smallest child (3) to fix heap order

Swap element 7 with its smallest child (6) to fix heap order

Both children of 7 are larger, stop swapping

Changing element priorities

First, find the position p of the element you want to change

(Simplest: keep a separate dictionary mapping elements to positions, update it whenever we change positions of elements)

To increase priority: fix heap at position p, same as for find-and-remove

To decrease priority, repeat:

1. If $p = 0$ or parent of p has smaller priority, return
2. Swap elements at positions p and parent
3. Set $p =$ parent and continue repeating

...swapping on a path upward from p instead of downward

For both increase and decrease, time is $O(\log n)$

Insert new element

To insert a new element x with priority p into a heap that already has n elements:

- Add x to array cell $H[n]$ with priority $+\infty$
- The heap property is satisfied!
  (because of the big but fake priority)
- Change the priority of x to p
  (swapping along path towards root)

Time: $O(\log n)$

Heapify

Given an unsorted array, put it into heap order

Main idea: Fix subtrees that have only one element out of order

(by swapping with children on downward path in tree)

Recursive version:

To recursively heapify the subtree rooted at i:

1. Recursively heapify left child of i
2. Recursively heapify right child of i
3. Fix heap at i

Non-recursive version:

For $i = n-1, n-2, \ldots 0$:
    Fix heap at i

Heapify analysis (either version)

In a binary heap with n tree nodes:

- $\lceil n/2 \rceil$ are leaves, fix heap takes one step
- $\lceil n/4 \rceil$ are parents of leaves, fix heap takes two steps

  ...
- $\lceil n/2^i \rceil$ are i levels up, fix heap takes i steps

Total steps:

$$\sum_{i=1}^{\lceil \log_2(n+1) \rceil} i \left\lceil \frac{n}{2^i} \right\rceil \leq O(\log^2 n) + n \sum_{i=1}^{\infty} \frac{i}{2^i} = 2n + O(\log^2 n)$$

The $O(\log^2 n)$ term is how much the rounding-up part of the formula can contribute, compared to not rounding.)

So total time is $O(n)$, as stated earlier

k-ary heaps

Can we speed this up?

In a comparison model of computation, some priority queue operations must take $\Omega(\log n)$ time

- We can use a priority queue to sort n items by repeatedly finding and removing the smallest one (heapsort)
- But comparison-based sorting requires $\Omega(n \log n)$ time

But in Dijkstra, different operations have different frequencies

- In a graph with n vertices and m edges...
- Dijkstra does n find-and-remove-min operations,
- but up to m decrease-priority operations
- In some graphs, m can be much larger than n

Goal: speed up the more frequent operation (decrease-priority) without hurting the other operations too much

k-ary heap

Same as binary heap but with $k > 2$ children per node

$\text{Parent}(x) = \lfloor (x-1)/k \rfloor$
$\text{Children}(x) = (kx+1, kx+2, \ldots kx+k)$

Height of tree: $\log_k n = \dfrac{\log n}{\log k}$

Time to decrease priority $= O(\text{height}) = O\left(\dfrac{\log n}{\log k}\right)$

(because path to root is shorter)

Time to find-and-remove-min $= O\left(\dfrac{k \log n}{\log k}\right)$

(because each step finds smallest of k children)

Dijkstra using k-ary heap

Time for m decrease-priority operations: $O\left(m\frac{\log n}{\log k}\right)$

Time for n find-and-remove-min operations: $O\left(nk\frac{\log n}{\log k}\right)$

To minimize total time, choose k to balance these two bounds

$$k = \max(2, \lceil m/n \rceil)$$

$$\text{Total time} = O\left(m\frac{\log n}{\log m/n}\right)$$

This becomes $O(m)$ whenever $m = \Omega(n^{1+\varepsilon})$ for any constant $\varepsilon > 0$

k-ary heaps in practice

4-ary heaps may be better than 2-ary heaps for all operations

For heaps that are too large to fit into cache memory, even larger choices of k may be better

(Lower tree height beats added complexity of more children, especially when each level costs a cache miss)

See `https://en.wikipedia.org/wiki/D-ary_heap` for references