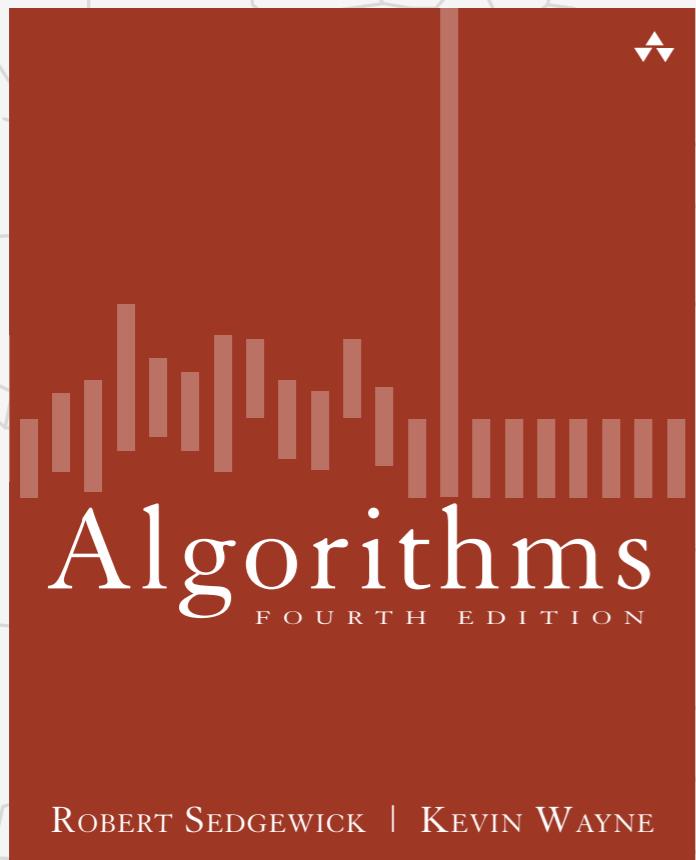


Algorithms

ROBERT SEDGEWICK | KEVIN WAYNE



<http://algs4.cs.princeton.edu>

2.4 PRIORITY QUEUES

- *API and elementary implementations*
- *binary heaps*
- *heapsort*
- *event-driven simulation*

BigData Scenario

- *What if you have a collection of N objects but you want to process the M biggest/best ones? You could use quickSelect from last week's lecture. Or you could sort everything and then take the first M elements, right?*
- *But what if the collection is too large? Or it's an infinite stream?*
- *Remember quickSelect and how we were able to get the k^{th} element in linear time?*
- *Remember all of those binary trees that we've thought about but didn't actually implement? Well, now we're going to start using them!*

Algorithms

ROBERT SEDGEWICK | KEVIN WAYNE

<http://algs4.cs.princeton.edu>

2.4 PRIORITY QUEUES

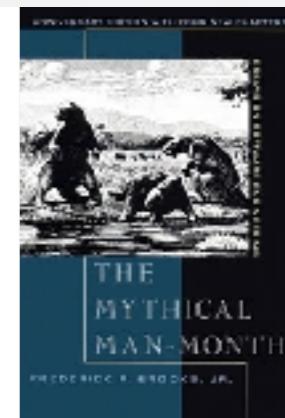
- *API and elementary implementations*
- *binary heaps*
- *heapsort*
- *event-driven simulation*

Collections

A **collection** is a data types that store groups of items.

data type	key operations	data structure
stack	PUSH, POP	<i>linked list, resizing array</i>
queue	ENQUEUE, DEQUEUE	<i>linked list, resizing array</i>
priority queue	INSERT, DELETE-MAX	<i>binary heap</i>
symbol table	PUT, GET, DELETE	<i>BST, hash table</i>
set	ADD, CONTAINS, DELETE	<i>BST, hash table</i>

“Show me your code and conceal your data structures, and I shall continue to be mystified. Show me your data structures, and I won’t usually need your code; it’ll be obvious.” — Fred Brooks



Priority queue

Collections. Insert and delete items. Which item to delete?

Stack. Remove the item most recently added.

Queue. Remove the item least recently added.

Randomized queue. Remove a random item.

Priority queue. Remove the largest (or smallest) item.

<i>operation</i>	<i>argument</i>	<i>return value</i>
<i>insert</i>	P	
<i>insert</i>	Q	
<i>insert</i>	E	
<i>remove max</i>		Q
<i>insert</i>	X	
<i>insert</i>	A	
<i>insert</i>	M	
<i>remove max</i>		X
<i>insert</i>	P	
<i>insert</i>	L	
<i>insert</i>	E	
<i>remove max</i>		P

Priority queue API

Requirement. Generic items are Comparable.

public class MaxPQ<Key extends Comparable<Key>>		Key must be Comparable (bounded type parameter)
MaxPQ()		<i>create an empty priority queue</i>
void insert(Key v)		<i>insert a key into the priority queue</i>
Key delMax()		<i>return and remove the largest key</i>
boolean isEmpty()		<i>is the priority queue empty?</i>
Key max()		<i>return the largest key</i>
int size()		<i>number of entries in the priority queue</i>

Priority queue applications

- Big data application. [fraud detection]
- Event-driven simulation. [customers in a line, colliding particles]
- Numerical computation. [reducing roundoff error]
- Data compression. [Huffman codes]
- Graph searching. [Dijkstra's algorithm, Prim's algorithm]
- Number theory. [sum of powers]
- Artificial intelligence. [A* search]
- Statistics. [online median in data stream]
- Operating systems. [load balancing, interrupt handling]
- Computer networks. [web cache]
- Discrete optimization. [bin packing, scheduling]
- Spam filtering. [Bayesian spam filter]

Generalizes: stack, queue, randomized queue.

Example: Election-related tweets in 2016

Table 1-1

user_key	created_str	text	hashtags	score
ryanmaxwell_1	2016-03-22 18:31:42	Kills Are you trying to say that there were no terrorist attacks in Europe before refugees were let in?		
detroitdailynew	2016-10-10 20:57:00	Clinton: Trump should've apologized more, attacked less https://t.co/eJampkoHFZ	[]	10
cookncooks	2017-02-22 12:43:43	RT @Itapoll: Who was/is the best president of the past 25 years? (Vote & Retweet)	[]	
queenofthewo	2016-12-26 15:06:41	RT @jww372: I don't have to guess your religion! #ChristmasAftermath	["ChristmasAfte"]	
mrclydepratt	2017-08-06 02:36:24	RT @Shareblue: Pence and his lawyers decided which of his official emails the public could see https://t.co/HjhPguBK1Y by @alisonrc	[]	
giselleevns	2016-10-26 15:33:58	@ModicaGiunta me, too!	[]	
baobaeham	2017-03-07 18:11:44	RT @MDBlanchfield: You'll never guess who tweeted something false that he saw on TV - The Washington Post https://t.co/K2e4XdXf	[]	
judelambertusa	2016-12-30 12:49:30	RT @100PercFEDUP: New post: WATCH: DIAMOND AND SILK Rip On John Kerry Over Israel Comments (VIDEO) https://t.co/NkdKaC	[]	
ameliebaldwin	2016-10-30 01:48:19	RT @AriaWilsonGOP: 3 Women Face Charges After Being Caught Stealing Dozens Of Trump Signs https://t.co/JjlZxaW3JN https://t.c	[]	5
hiimkhloe	2016-03-16 19:07:39	One of the ways to remind that #BlackLivesMatter #BlackPressDay	["Blacklivesmat"]	
jasper_fly	2017-01-11 15:46:49	RT @EIPinguinito: #myfarewellwordswouldbe I've buried my fortune in the park under a giant...	["MyFarewellWc"]	
patriotblake	2016-12-12 11:43:11	RT @America_1st_: CW: "The thing that impressed me was that Trump is always comfortable in own skin, but now he was comfortable	[]	5
anthonywoodboy	2015-06-10 20:33:31	RT @AllAmericanGirl: 🇺🇸Obama to add 450 Iraq military advisers http://t.co/nNAeycMv6c	[]	5
pamela_moore13	2016-08-31 01:49:22	Dave Chappelle: "Black Lives Matter" is the worst slogan I've ever heard! How about "enough is enough"? VotingTrump! https://t.co/5	[]	
laurabaeley	2017-01-23 08:59:25	RT @RadioACR: Chuck Todd vs Kellyanne Conway... a contentious interview here, @21WIRE says @KellyannePolls owned Chuck Tod	[]	
giselleevns	2017-01-04 15:32:27	#My2017BiggestHope to reach this level of pettiness https://t.co/YchECGXOTI	["My2017Bigge"]	
pamela_moore13	2016-09-19 23:11:43	The war is here! This gentleman made more sense in 30 sec than #Obama for all time of his presidency.. https://t.co/OHueWuqwhO	["Obama"]	10
brianaregland	2017-01-11 10:36:18	RT @HipHopDX: .@Migos Scores FIRST #1 Hot 100 Hit With "Bad And Boujee" https://t.co/TM77Nbdyed https://t.co/HDhyOp7TE	["1"]	
patriotraphael	2015-07-13 01:57:20	RT @LibertyBritt: He's the brilliant guy who shoots himself in the foot to spite his face. And tries to convince us to do it too. https://...	[]	
hyddrox	2017-02-09 10:26:17	RT @K1erry: The Marco Rubio knockdown of Elizabeth Warren no liberal media outlet will cover https://t.co/Rh391fEXe3	[]	8
kansasdailynews	2016-07-27 14:55:39	Obama on Trump winning: 'Anything's possible' https://t.co/MjVMZ5TR8Y #politics	["Politics"]	19

Priority queue client example

Challenge. Find the largest M items in a stream of N items.

- Fraud detection: isolate \$\$ transactions.
- NSA monitoring: flag most suspicious documents.

N huge, M large

Constraint. Not enough memory to store N items.

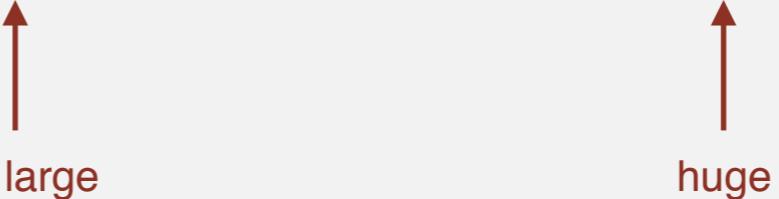
```
MaxPQ<Transaction> pq = new MaxPQ<Transaction>(M);  
  
while (StdIn.hasNextLine())  
{  
    String line = StdIn.readLine();  
    Transaction item = new Transaction(line);  
    pq.insert(item);  
}
```

pq contains
largest M items

Transaction data
type is Comparable
(ordered by \$\$)

Priority queue client example

Challenge. Find the largest M items in a stream of N items.



implementation	time	space
sort	$N \log N$	N
elementary PQ	$N M$	M
heap-based PQ	$N \log M$	M
best in theory	N	M

order of growth of finding the largest M in a stream of N items

Priority queue: unordered and ordered array implementation

<i>operation</i>	<i>argument</i>	<i>return value</i>	<i>size</i>	<i>contents (unordered)</i>	<i>contents (ordered)</i>
<i>insert</i>	P		1	P	P
<i>insert</i>	Q		2	P Q	P Q
<i>insert</i>	E		3	P Q E	E P Q
<i>remove max</i>		Q	2	P E	E P
<i>insert</i>	X		3	P E X	E P X
<i>insert</i>	A		4	P E X A	A E P X
<i>insert</i>	M		5	P E X A M	A E M P X
<i>remove max</i>		X	4	P E M A	A E M P
<i>insert</i>	P		5	P E M A P	A E M P P
<i>insert</i>	L		6	P E M A P L	A E L M P P
<i>insert</i>	E		7	P E M A P L E	A E E L M P P
<i>remove max</i>		P	6	E M A P L E	A E E L M P P

A sequence of operations on a priority queue

Priority queue: unordered array implementation

```
public class UnorderedArrayMaxPQ<Key extends Comparable<Key>>
{
    private final Key[] pq; // pq[i] = ith element on pq
    private int M = 0; // number of elements on pq

    public UnorderedArrayMaxPQ(int capacity)
    { pq = (Key[]) new Comparable[capacity]; }

    public boolean isEmpty()
    { return M == 0; }

    public void insert(Key x)
    { pq[M++] = x; }

    public Key delMax()
    {
        int max = 0;
        for (int i = 1; i < M; i++)
            if (less(max, i)) max = i;
        swap(max, M-1);
        return pq[−M];
    }
}
```

This method is called M times

less() and swap()
similar to sorting methods
(but don't pass pq[])

should null out entry
to prevent loitering

Priority queue elementary implementations

Challenge. Implement **all** operations efficiently.

Note: we can minimize inadvertent losses of elements by making the priority queue bigger than M (by factor of c).

implementation	insert	del max	Total ops
unordered array	1	cM	$N + cM^2$
ordered array	cM	1	$M(cN+1)$
goal	$\lg cM$	$\lg cM$	$(N+M) \lg cM$

**order of growth of running time for priority queue of size cM
with N inserts and M delMax operations; c is the “safety” factor.**

Algorithms

ROBERT SEDGEWICK | KEVIN WAYNE

<http://algs4.cs.princeton.edu>

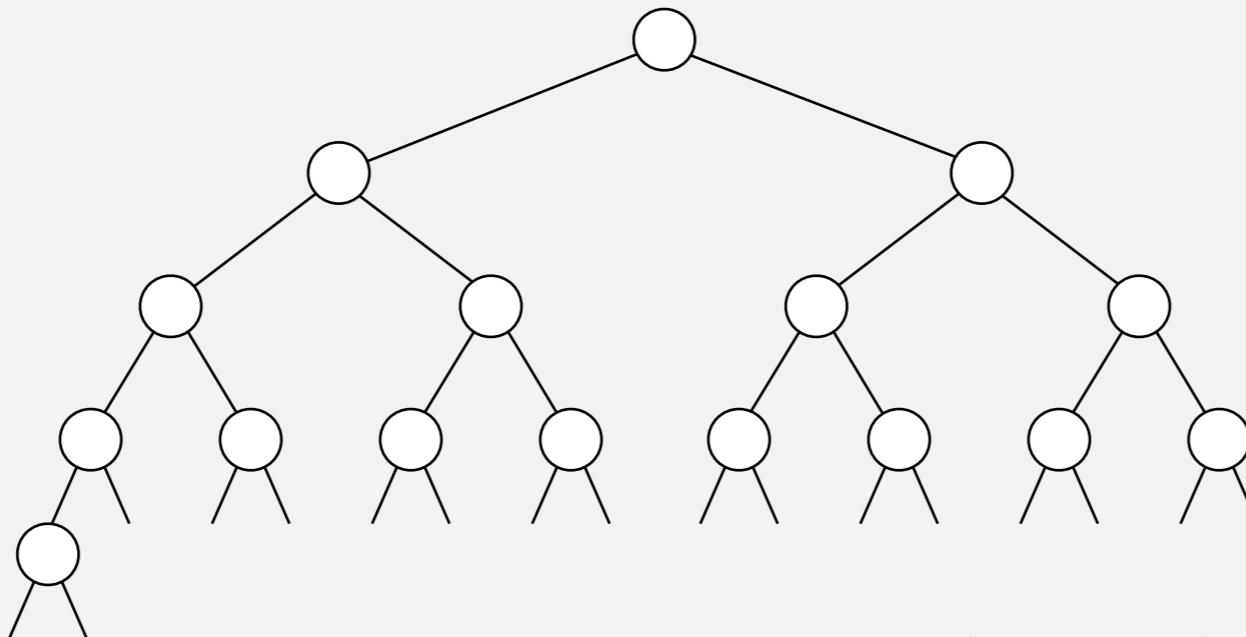
2.4 PRIORITY QUEUES

- *API and elementary implementations*
- ***binary heaps***
- *heapsort*
- *event-driven simulation*

Complete binary tree

Binary tree. Empty or node with links to left and right binary trees.

Complete tree. Perfectly balanced, except for bottom level.



Property. Height of complete tree with N nodes is $\lfloor \lg N \rfloor$.

Pf. Height increases only when N is a power of 2.

A complete binary tree in nature



Hyphaene Compressa - Doum Palm

© Shlomit Pinter

Definitions

- Implicit:
 - *positional*, that's to say there are no (explicit) pointers—the relationships between elements are based on their relative positions. An *implicit* structure can be implemented using an *array*.
- In-place/efficient:
 - the element stored *is* or *yields* its own key (there is no separate *value* to be stored)
- Binary Tree:
 - A data structure in which each node has at most two children
- Complete Tree:
 - A tree in which every level is full except possibly the lowest (furthest from root) level
- Binary Heap:
 - A binary tree which is implemented as an in-place *implicit* data structure.
- Min Tree/Heap:
 - A tree such that the root is the *smallest* element, as opposed to a max tree/heap where the root is the *largest* element.

Binary heap representations

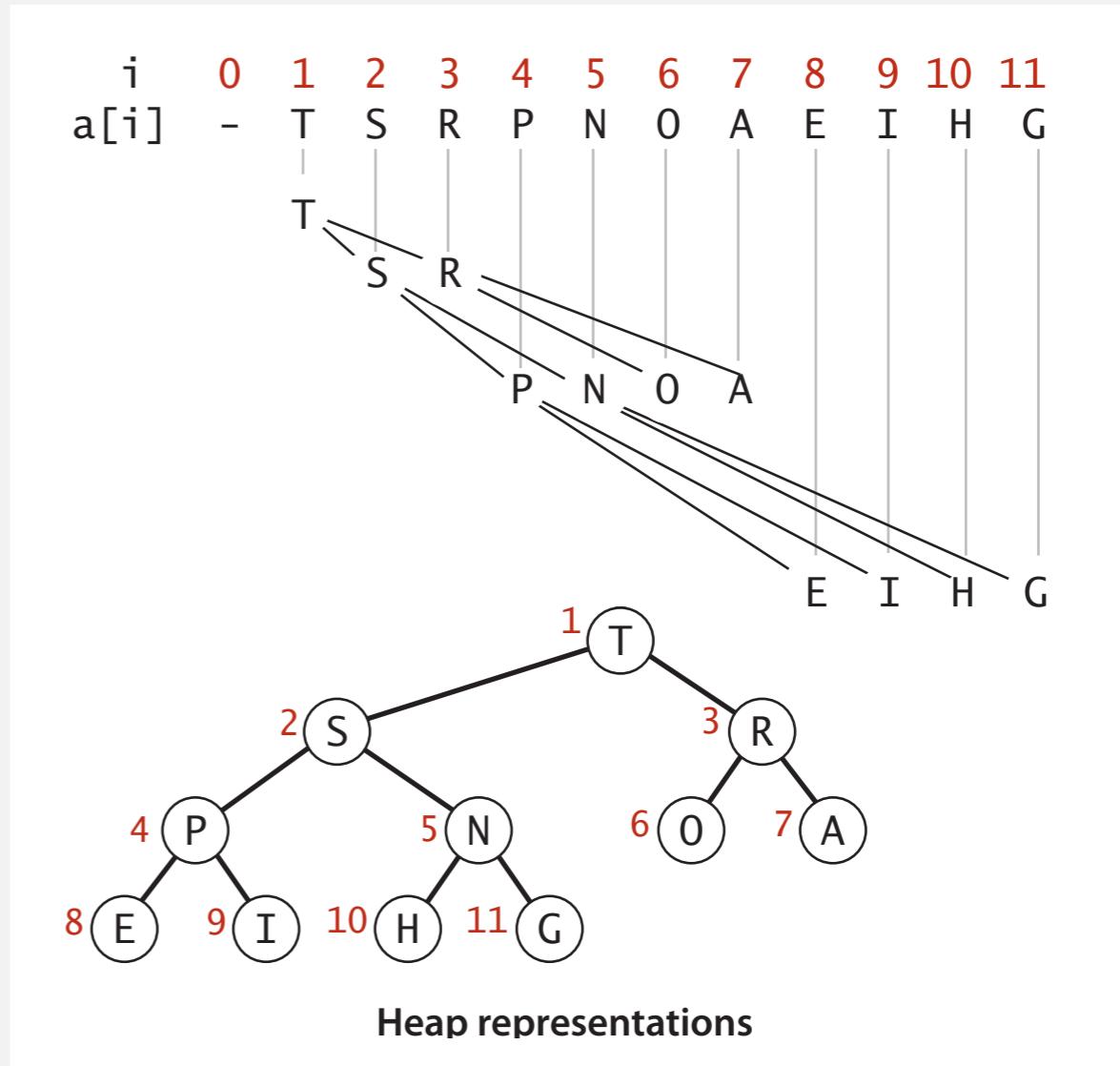
Binary heap. Array representation of a heap-ordered complete binary tree.

Heap-ordered binary tree.

- Keys in nodes.
- Parent's key no smaller than children's keys.
- Elements of levels are in *arbitrary* order.

Array representation.

- Indices start at 1.
- Take nodes in **level** order.
- No explicit links needed!

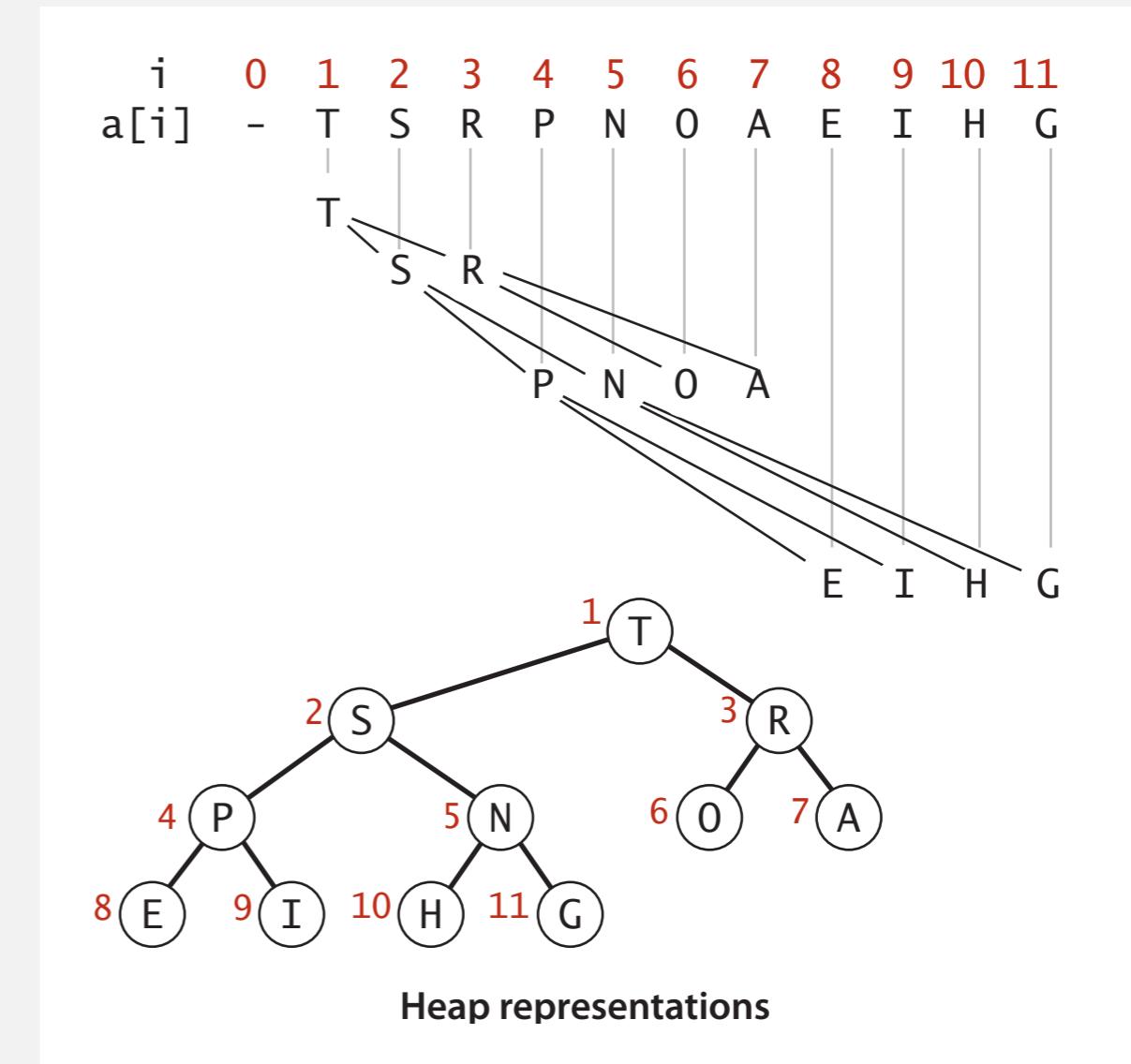


Binary heap properties

Proposition. Largest key is $a[1]$, which is root of binary tree.

Proposition. Can use array indices to move through tree.

- Parent of node at k is at $k/2$.
- Children of node at k are at $2k$ and $2k+1$.

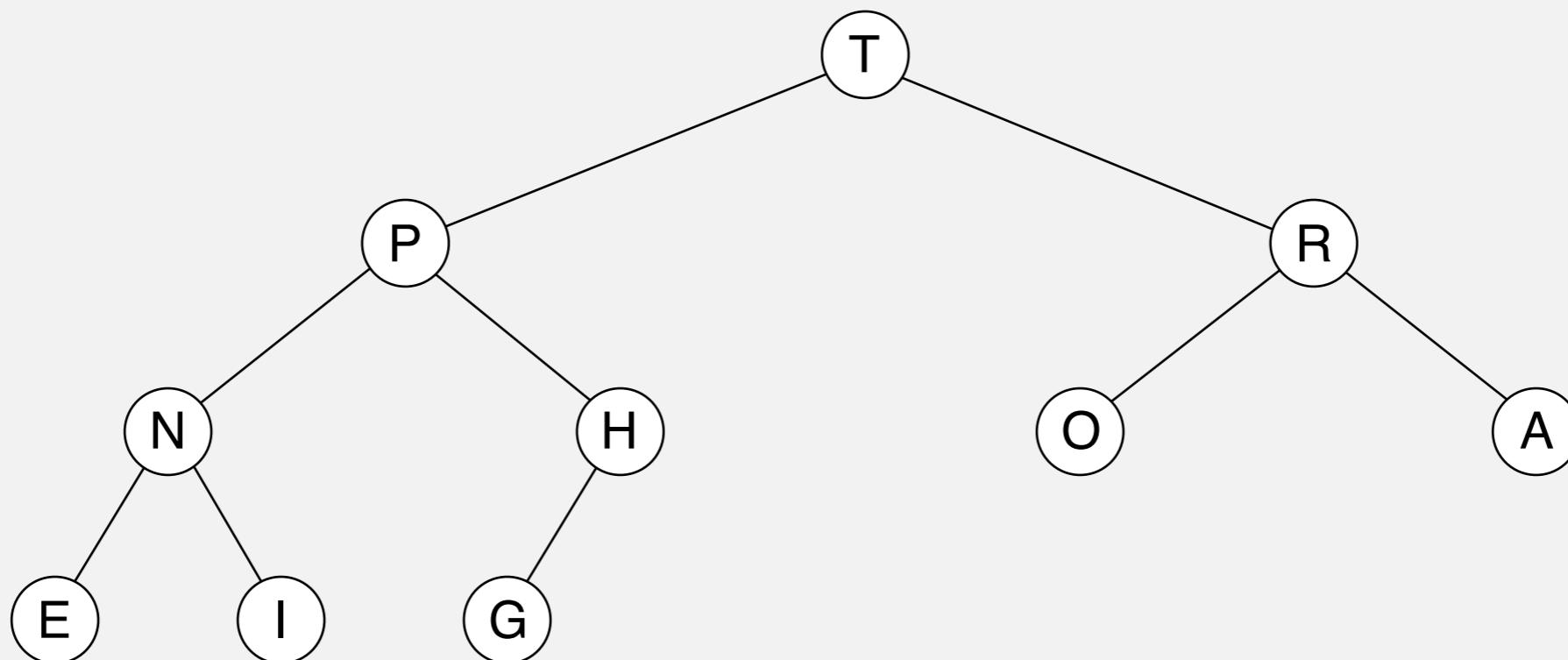


Binary heap demo

Insert. Add node at end, then swim it up.

Remove the maximum. Exchange root with node at end, then sink it down.

Heap is ordered



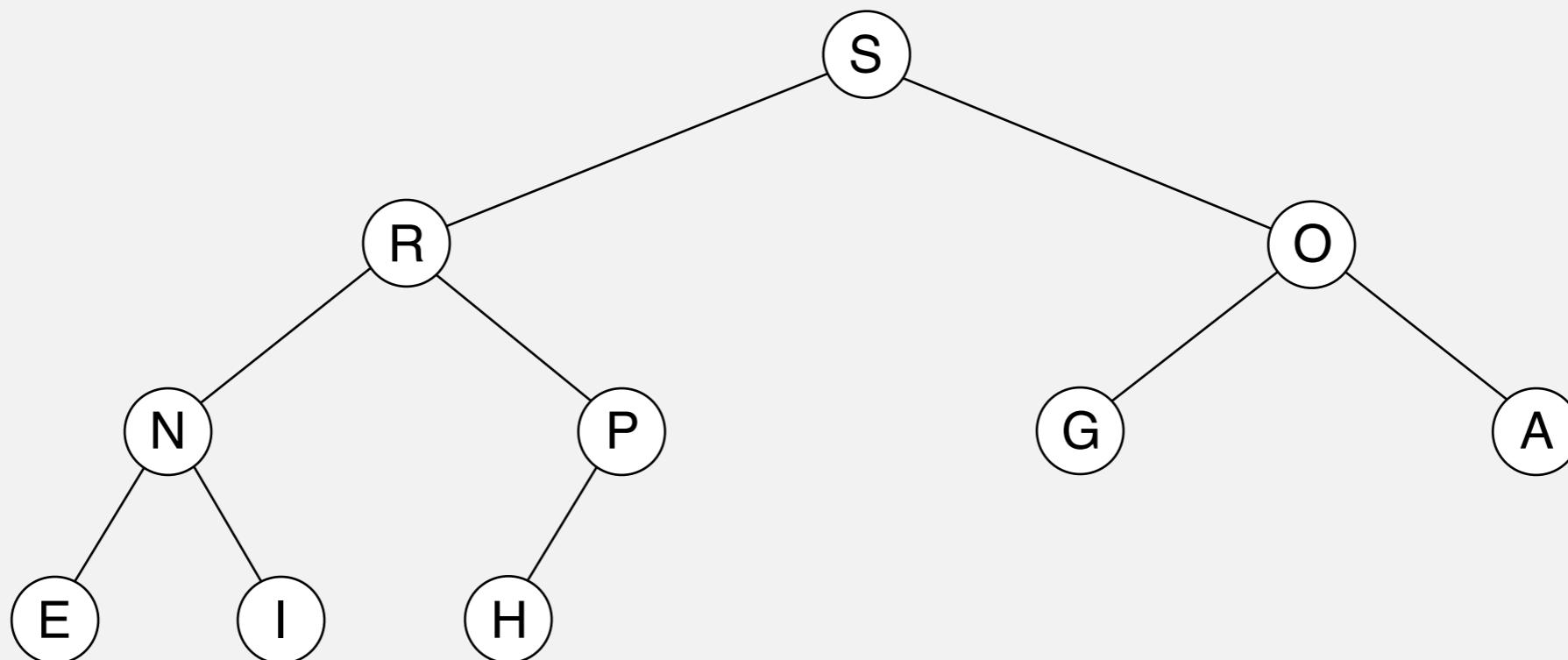
T	P	R	N	H	O	A	E	I	G	
---	---	---	---	---	---	---	---	---	---	--

Binary heap demo

Insert. Add node at end, then swim it up.

Remove the maximum. Exchange root with node at end, then sink it down.

Heap is ordered



Promotion in a heap

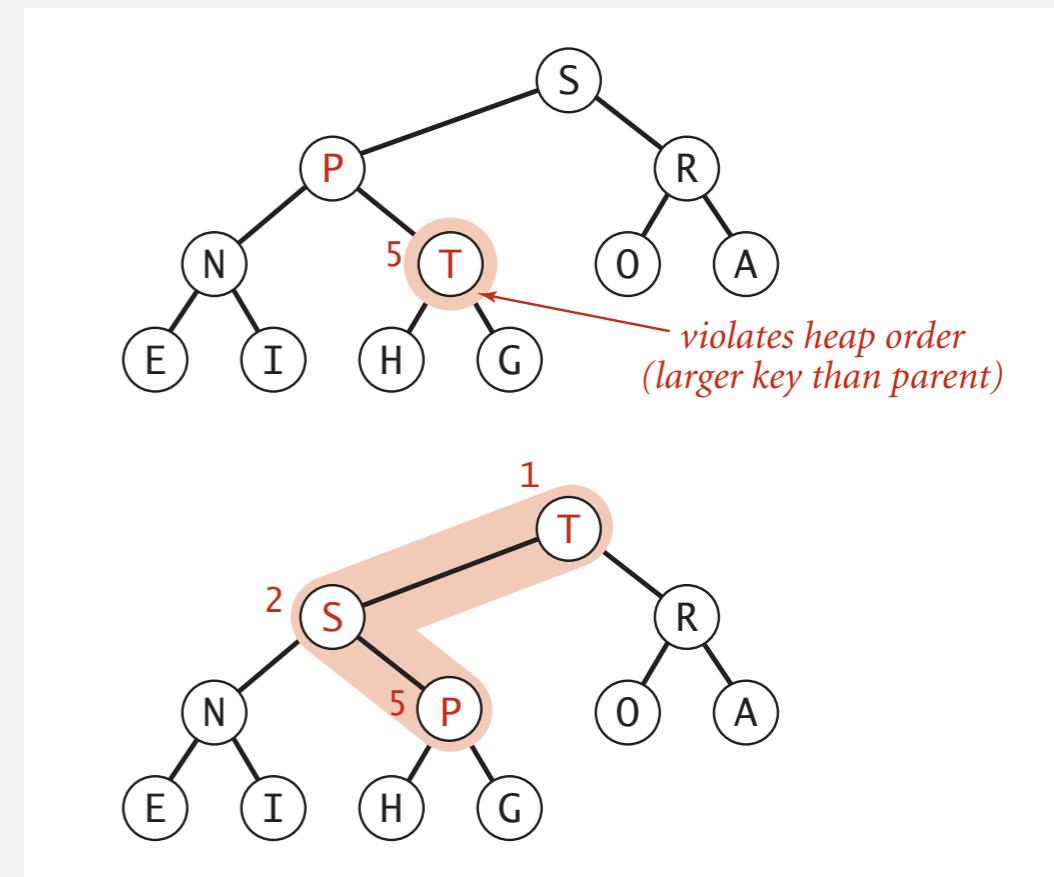
Scenario. Child's key becomes **larger** key than its parent's key.

To eliminate the violation:

- Exchange key in child with key in parent.
- Repeat until heap order restored.

```
private void swim(int k)
{
    while (k > 1 && less(k/2, k))
    {
        swap(k, k/2);
        k = k/2;
    }
}
```

parent of node at k is at k/2



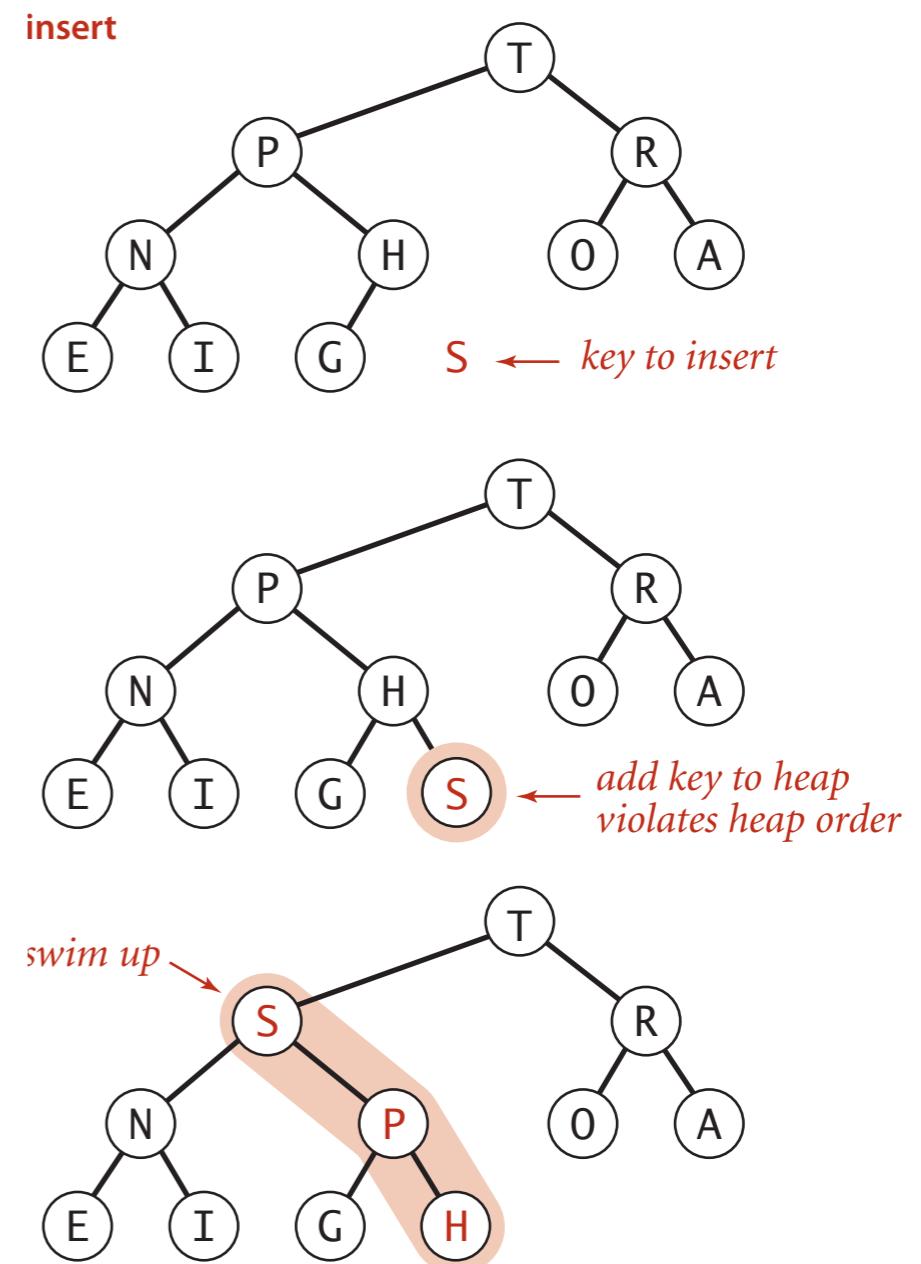
Insertion in a heap

Insert. Add node at end, then swim it up.

Cost. At most $1 + \lg M$ compares.

```
public void insert(Key x)
{
    if (M == pq.length) M--;
    pq[++M] = x;
    swim(M);
}
```

Ensure room,
losing last element
If necessary



Demotion in a heap

Scenario. Parent's key becomes **smaller** than one (or both) of its children's.

To eliminate the violation:

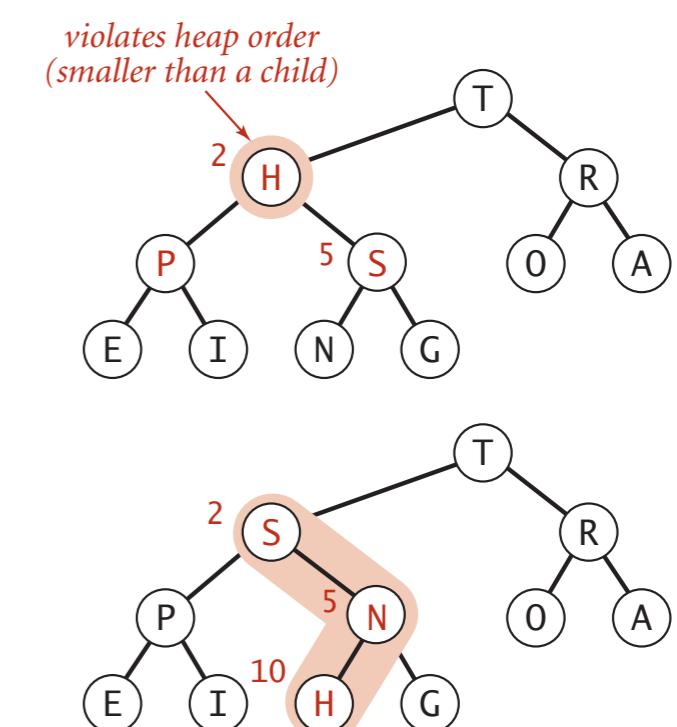
- Exchange key in parent with key in larger child.
- Repeat until heap order restored.

why not smaller child?

```
private void sink(int k)
{
    while (2*k <= M)
    {
        int j = 2*k;
        if (j < M && less(j, j+1)) j++;
        if (!less(k, j)) break;
        swap(k, j);
        k = j;
    }
}
```

children of node at k are

2k and 2k+1

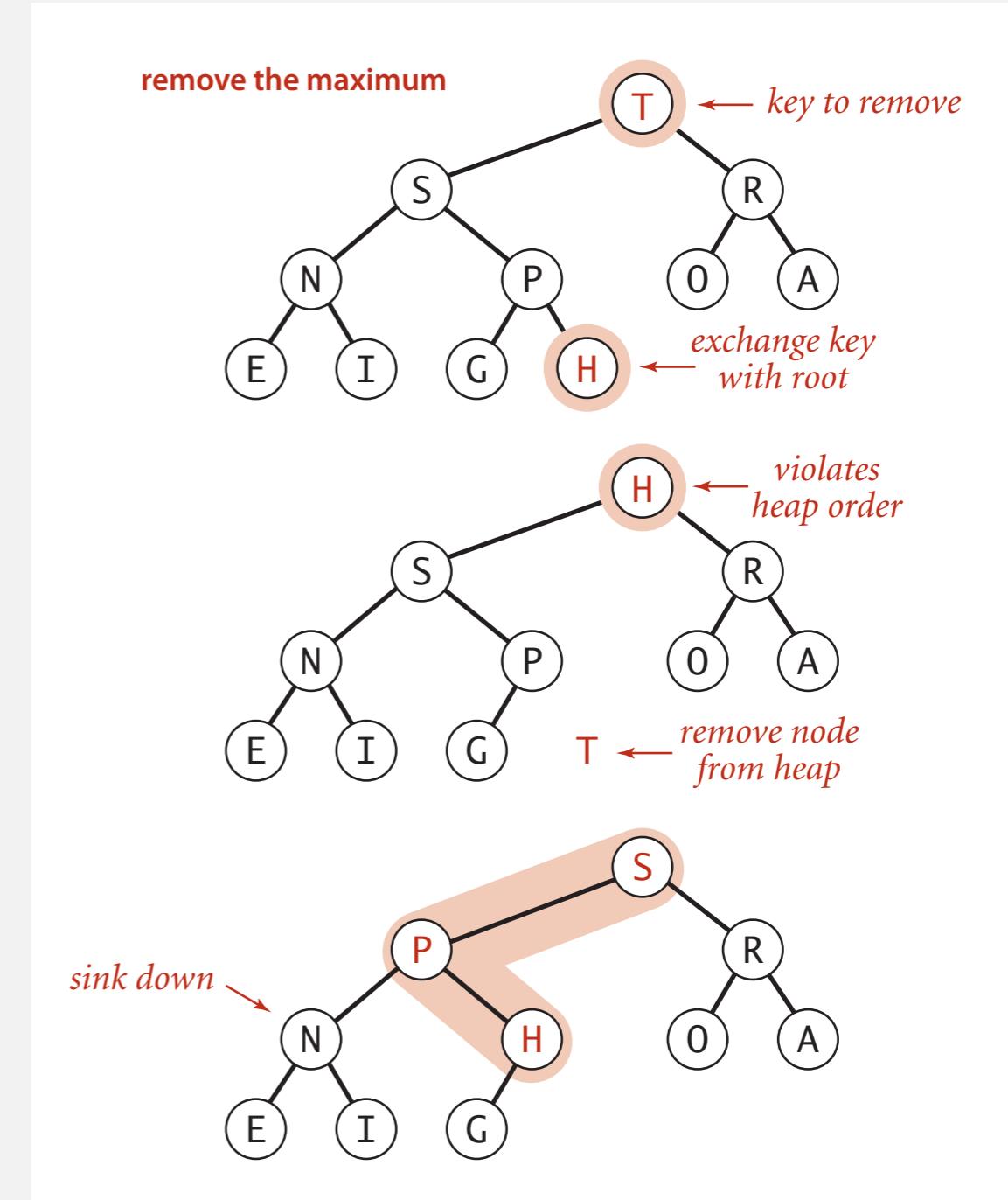


Delete the maximum in a heap

Delete max. Exchange root with node at end, then sink it down.

Cost. At most $2 \lg M$ compares.

```
public Key delMax()
{
    Key max = pq[1];
    swap(1, M--);
    sink(1);           ← prevent loitering
    pq[M+1] = null;
    return max;
}
```



Binary heap: Java implementation

```
public class MaxPQ<Key extends Comparable<Key>>
{
    private Key[] pq;
    private int M;

    public MaxPQ(int capacity)
    { pq = (Key[]) new Comparable[capacity+1]; }

    public boolean isEmpty()
    { return M == 0; }
    public void insert(Key key)
    public Key delMax()
    { /* see previous code */ }

    private void swim(int k)
    private void sink(int k)
    { /* see previous code */ }

    private boolean less(int i, int j)
    { return pq[i].compareTo(pq[j]) < 0; }
    private void swap(int i, int j)
    { Key t = pq[i]; pq[i] = pq[j]; pq[j] = t; }

}
```

fixed capacity
(for simplicity)

PQ ops

heap helper functions

array helper functions

Priority queues implementation cost summary

implementation	insert	del max	Total ops
unordered array	1	cM	$N + cM^2$
ordered array	cM	1	$M(cN+1)$
binary heap	$\lg cM$	$2 \lg cM$	$(N+2M) \lg cM$

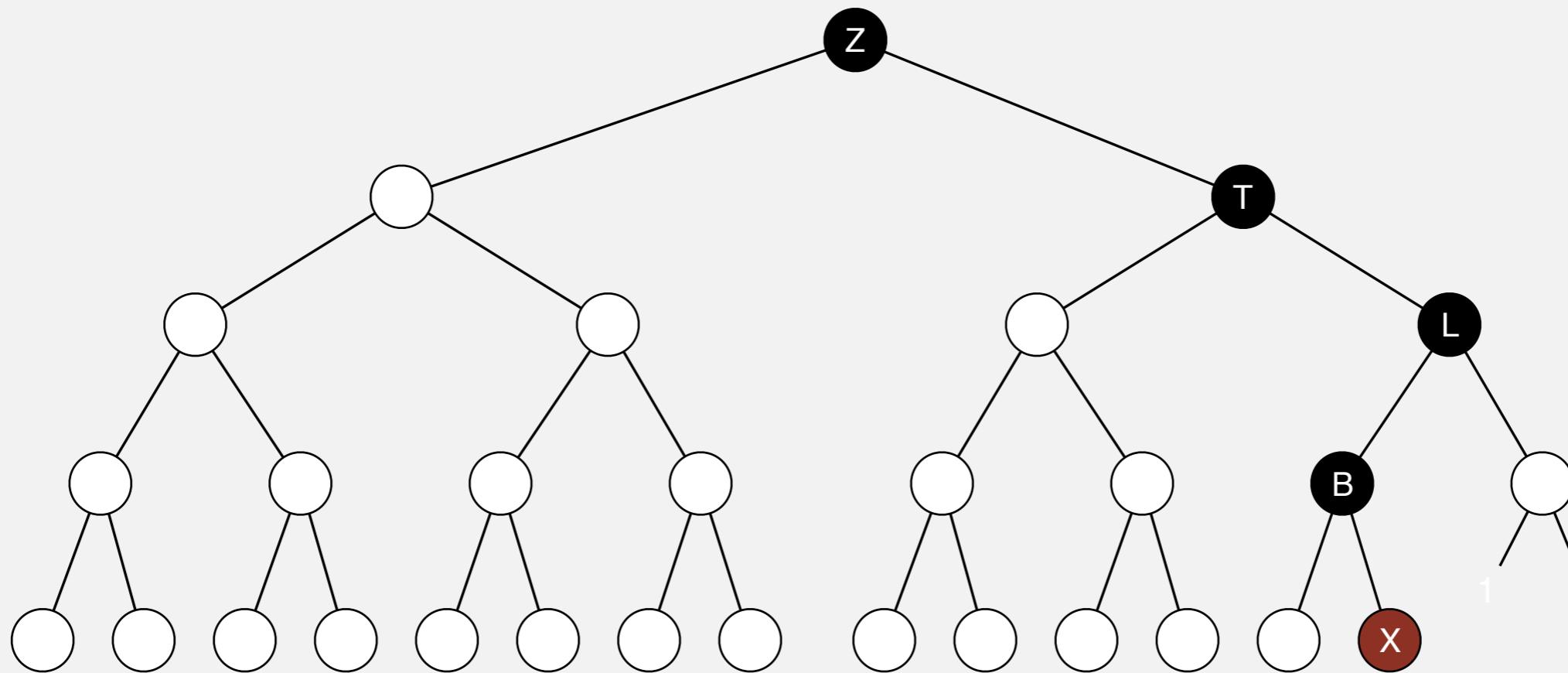
- ▶ “sink down” takes two compares per level whereas “swim up” requires only one per level.

order of growth of running time for priority queue of size cM with N inserts and M delMax operations; c is the “safety” factor of the priority queue.

Binary heap: practical improvements

Half-exchanges in sink and swim.

- Reduces number of array accesses.
- Worth doing.



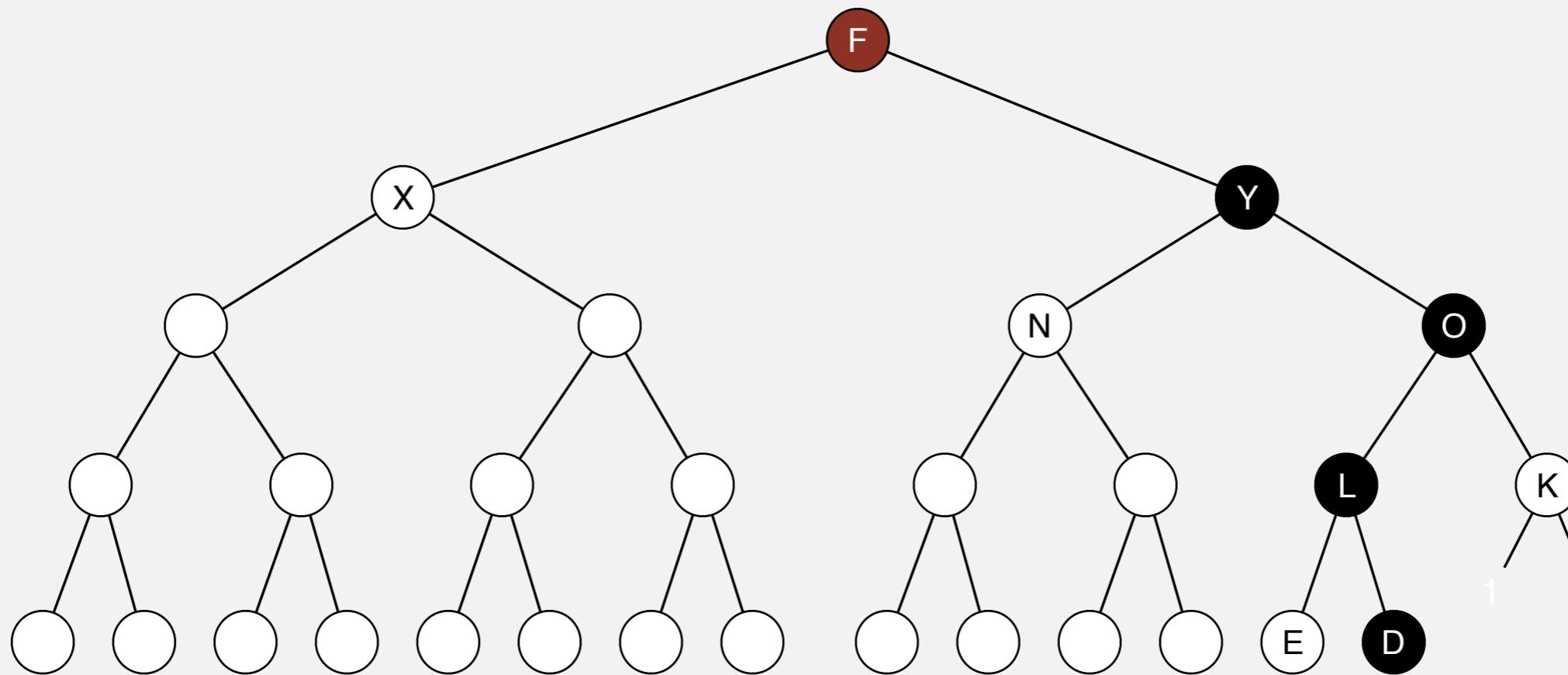
Binary heap: practical improvements

Floyd's sink-to-bottom trick.

- Sink key at root all the way to bottom. ← 1 compare per node
- Swim key back up. ← some extra compares and exchanges
- Fewer compares; more exchanges.
- Worthwhile depending on cost of compare and exchange.



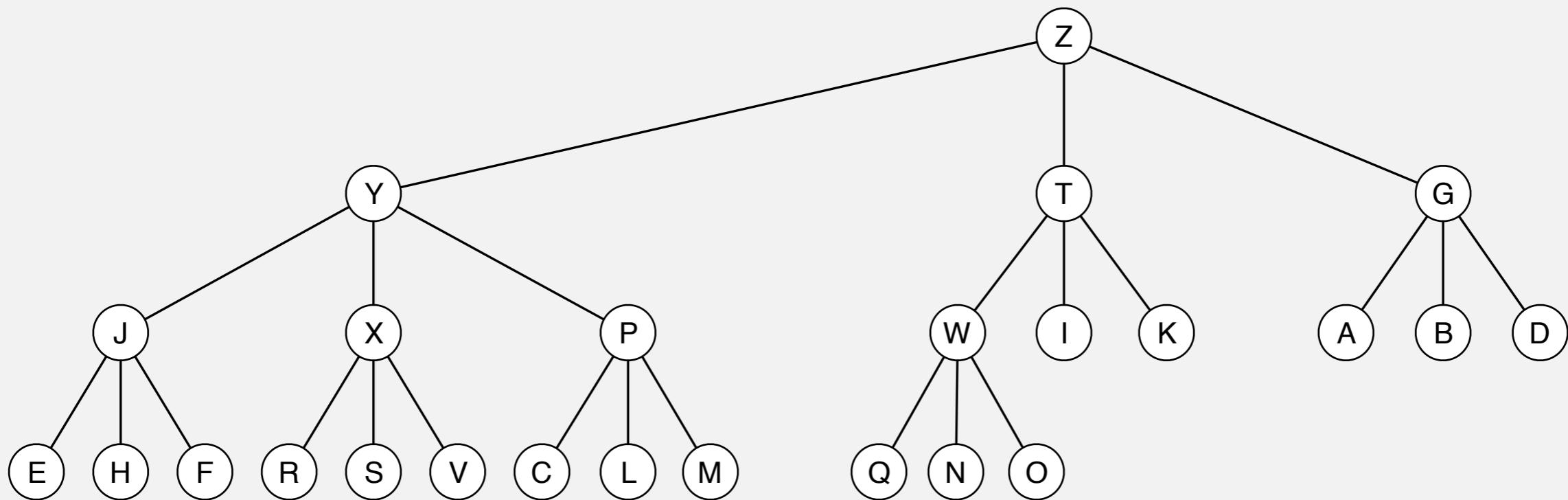
R. W. Floyd
1978 Turing award



Binary heap: practical improvements

Multiway heaps.

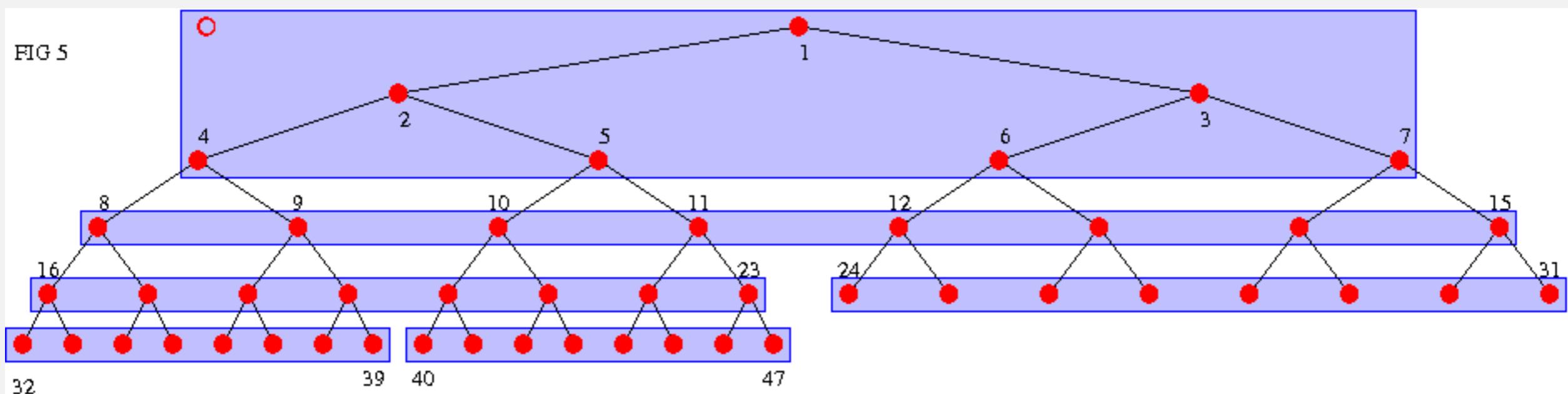
- Complete d -way tree.
- Parent's key no smaller than its children's keys.
- Swim takes $\log_d M$ compares; sink takes $d \log_d M$ compares.
- Sweet spot: $d = 4$.



3-way heap

Binary heap: practical improvements

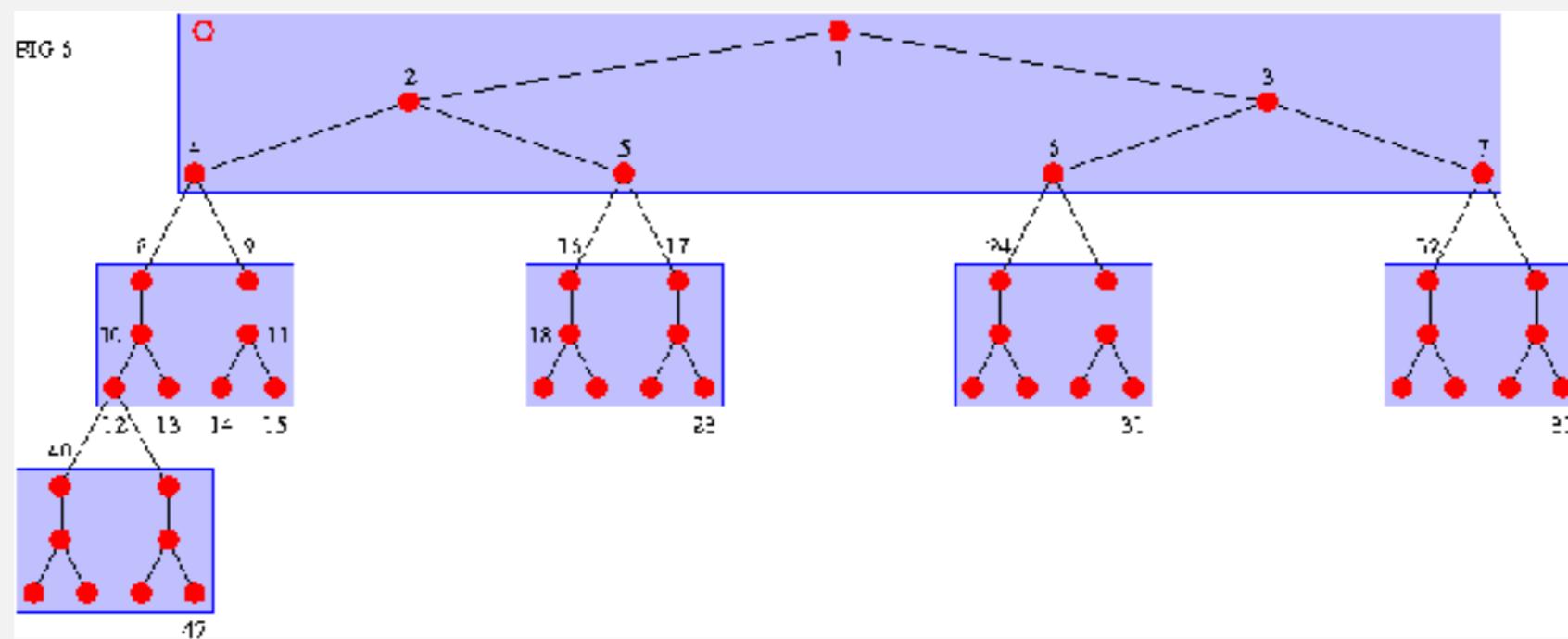
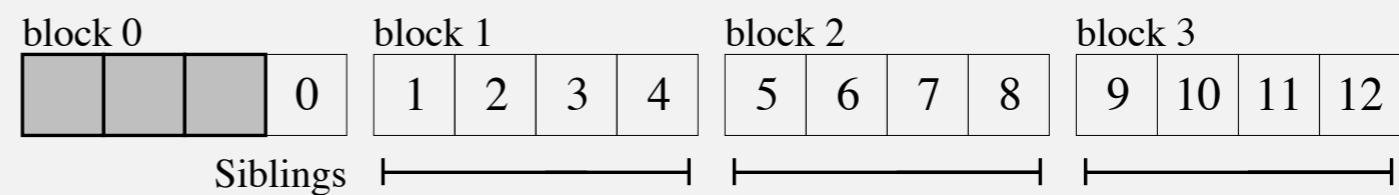
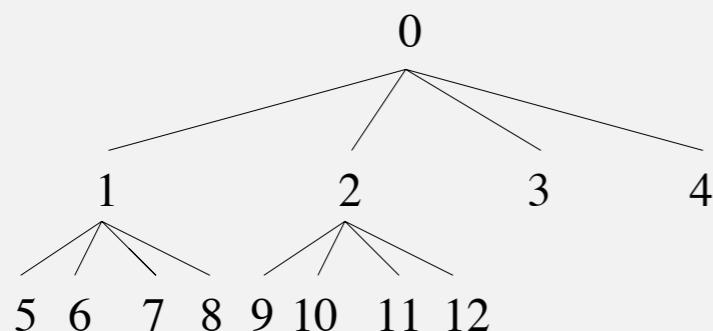
Caching. Binary heap is not cache friendly.



Binary heap: practical improvements

Caching. Binary heap is not cache friendly.

- Cache-aligned d -heap.
- Funnel heap.
- B-heap.
- ...



Priority queues implementation cost summary

implementation	insert	del max	max
unordered array	1	M	M
ordered array	M	1	1
binary heap	$\lg M$	$\lg M$	1
d-ary heap	$\log_d M$	$d \log_d M$	1
Fibonacci	1	$\log M^\dagger$	1
Brodal queue	1	$\log M$	1
impossible	1	1	1

← why impossible?

† amortized

order-of-growth of running time for priority queue with M items

Binary heap considerations

Underflow and overflow.

- Underflow: throw exception if deleting from empty PQ.
- Overflow: add no-arg constructor and use resizing array.

leads to log M
amortized time per op

Minimum-oriented priority queue.

- Replace less() with greater().
- Implement greater().

Other operations.

- Remove an arbitrary item.
- Change the priority of an item.

can implement efficiently with sink() and swim()
[stay tuned for Prim/Dijkstra]

Immutability of keys.

- Assumption: client does not change keys while they're on the PQ.
- Best practice: use immutable keys.

Immutability: implementing in Java

Data type. Set of values and operations on those values.

Immutable data type. Can't change the data type value once created.

```
public final class Vector {  
    private final int M;  
    private final double[] data;  
  
    public Vector(double[] data) {  
        this.M = data.length;  
        this.data = new double[M];  
        for (int i = 0; i < M; i++)  
            this.data[i] = data[i];  
    }  
    ...  
}
```

- ← can't override instance methods
- ← instance variables private and final
- ← defensive copy of mutable instance variables
- ← instance methods don't change instance variables

Immutable. String, Integer, Double, Color, Vector, Transaction, Point2D.

Mutable. StringBuilder, Stack, Counter, Java array.

Immutability: properties

Data type. Set of values and operations on those values.

Immutable data type. Can't change the data type value once created.

Advantages.

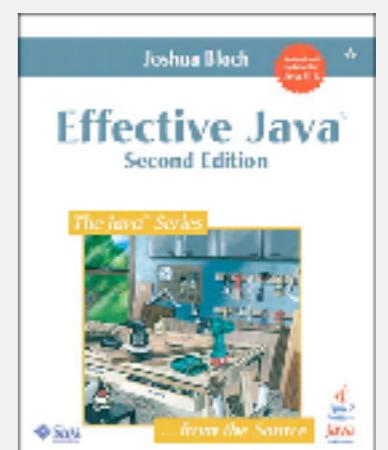
- Simplifies debugging.
- Safer in presence of hostile code.
- Simplifies concurrent programming.
- Safe to use as key in priority queue or symbol table.



Disadvantage. Must create new object for each data type value.

“Classes should be immutable unless there's a very good reason to make them mutable.... If a class cannot be made immutable, you should still limit its mutability as much as possible.”

— Joshua Bloch (Java architect)



Algorithms

ROBERT SEDGEWICK | KEVIN WAYNE

<http://algs4.cs.princeton.edu>

2.4 PRIORITY QUEUES

- *API and elementary implementations*
- *binary heaps*
- *heapsort*
- *event-driven simulation*

Sorting with a binary heap

Q. What is this sorting algorithm?

```
public void sort(String[] a)
{
    int M = a.length;
    MaxPQ<String> pq = new MaxPQ<String>();
    for (int i = 0; i < M; i++)
        pq.insert(a[i]);
    for (int i = M-1; i >= 0; i--)
        a[i] = pq.delMax();
}
```

Q. What are its properties?

A. $M \log M$, extra array of length M , not stable. What use could it be?

Intuition: Given that a heap is an array, can you think of any improvements?

- do sort in place.

How can we do this sort in-place?

- If we insist on creating the MaxPQ as in the previous algorithm, we cannot escape using twice the memory;
- But what if we could treat the (mutable) array that we are given as a binary heap?
- In order to fix the heap order, we will use the lower-level methods: *swim* and *sink*.

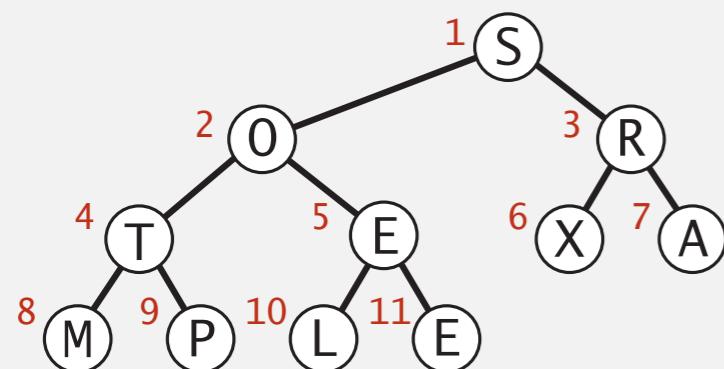
Heapsort

Basic plan for in-place sort.

- View input array as a complete binary tree.
- Heap construction: build a max-heap with all M keys.
- Sortdown: repeatedly remove the maximum key. ← But don't null it out!

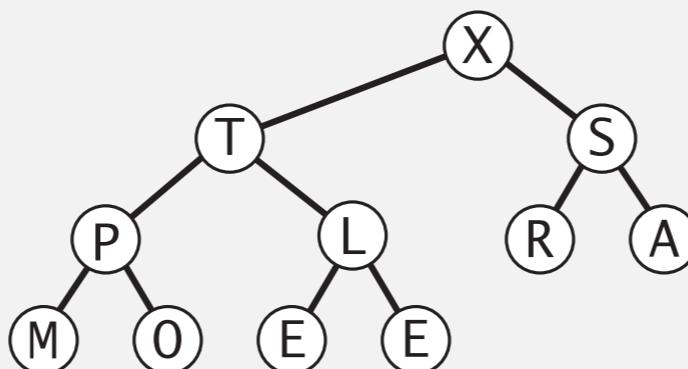
Heap order but not sort order

keys in arbitrary order



1	2	3	4	5	6	7	8	9	10	11
S	O	R	T	E	X	A	M	P	L	E

build max heap
(in place)



1	2	3	4	5	6	7	8	9	10	11
X	T	S	P	L	R	A	M	O	E	E

sorted result
(in place)



1	2	3	4	5	6	7	8	9	10	11
A	E	E	L	M	O	P	R	S	T	X

How to construct the heap?

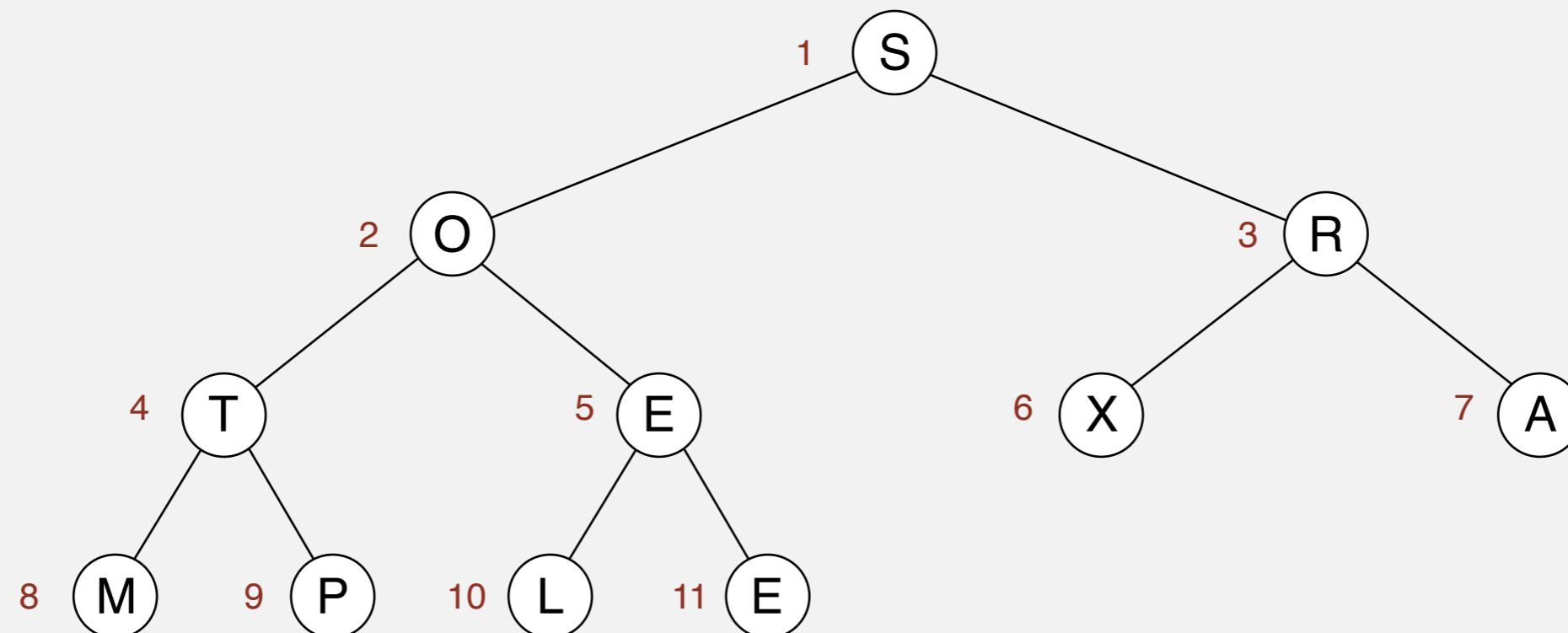
- Loop through the elements from the **last to the root** and for each element “fix” the heap order of the sub-heap for which that element is the root.
- We fix each sub-heap by applying *sink down* to the element.
- In practice, we can skip the first $N/2$ elements we encounter because each of these corresponds to single-element heap which is in heap order by definition.

Heapsort demo

Heap construction. Build max heap using bottom-up method.

we assume array entries are indexed 1 to N

array in arbitrary order

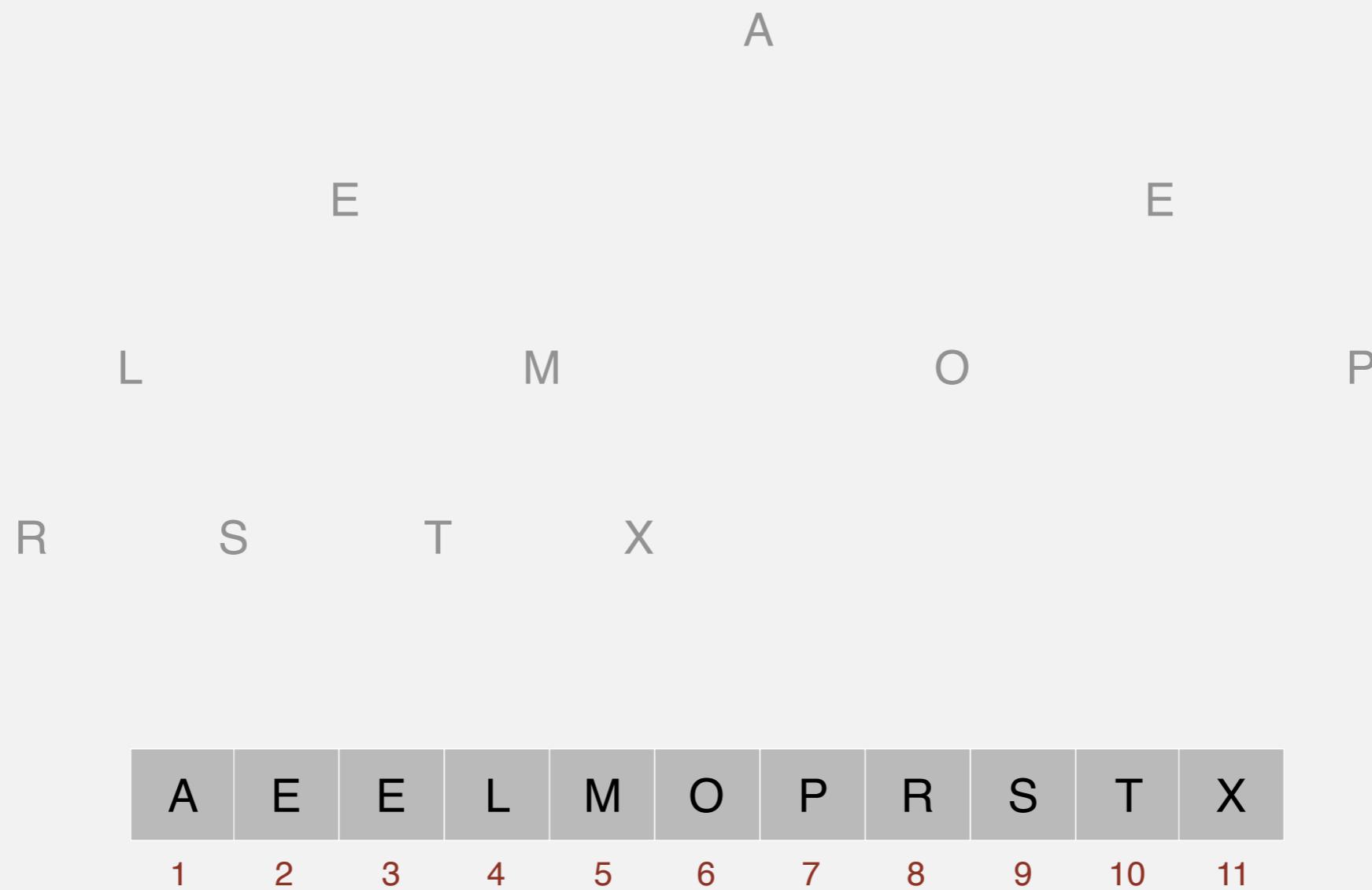


S	O	R	T	E	X	A	M	P	L	E
1	2	3	4	5	6	7	8	9	10	11

Heapsort demo

Sortdown. Repeatedly delete the largest remaining item.

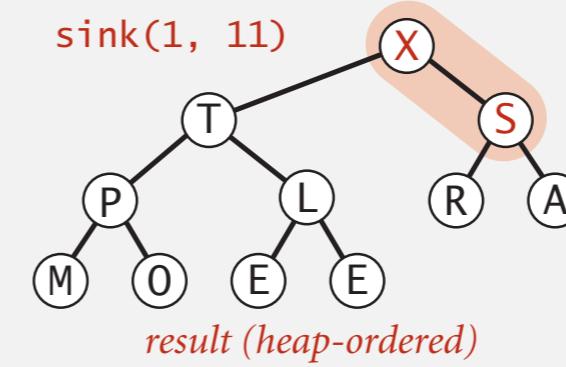
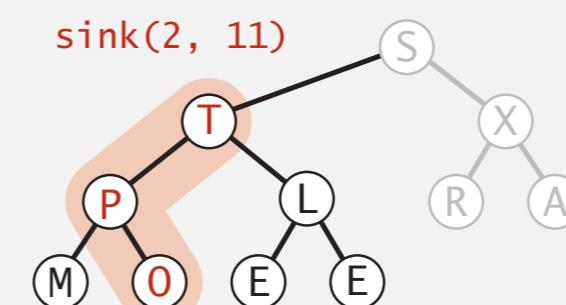
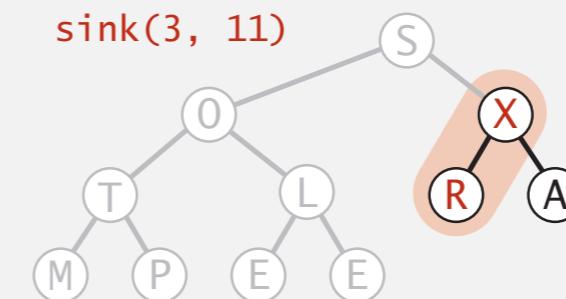
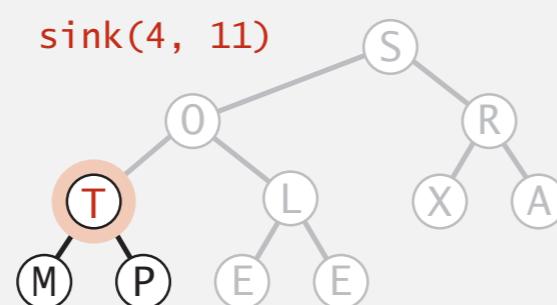
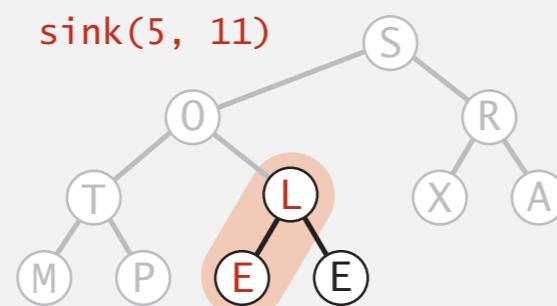
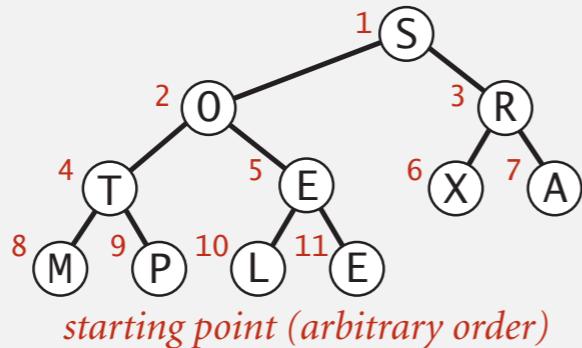
array in sorted order



Heapsort: heap construction

First pass. Build heap using bottom-up method.

```
for (int k = N/2; k >= 1; k--)  
    sink(a, k, N);
```

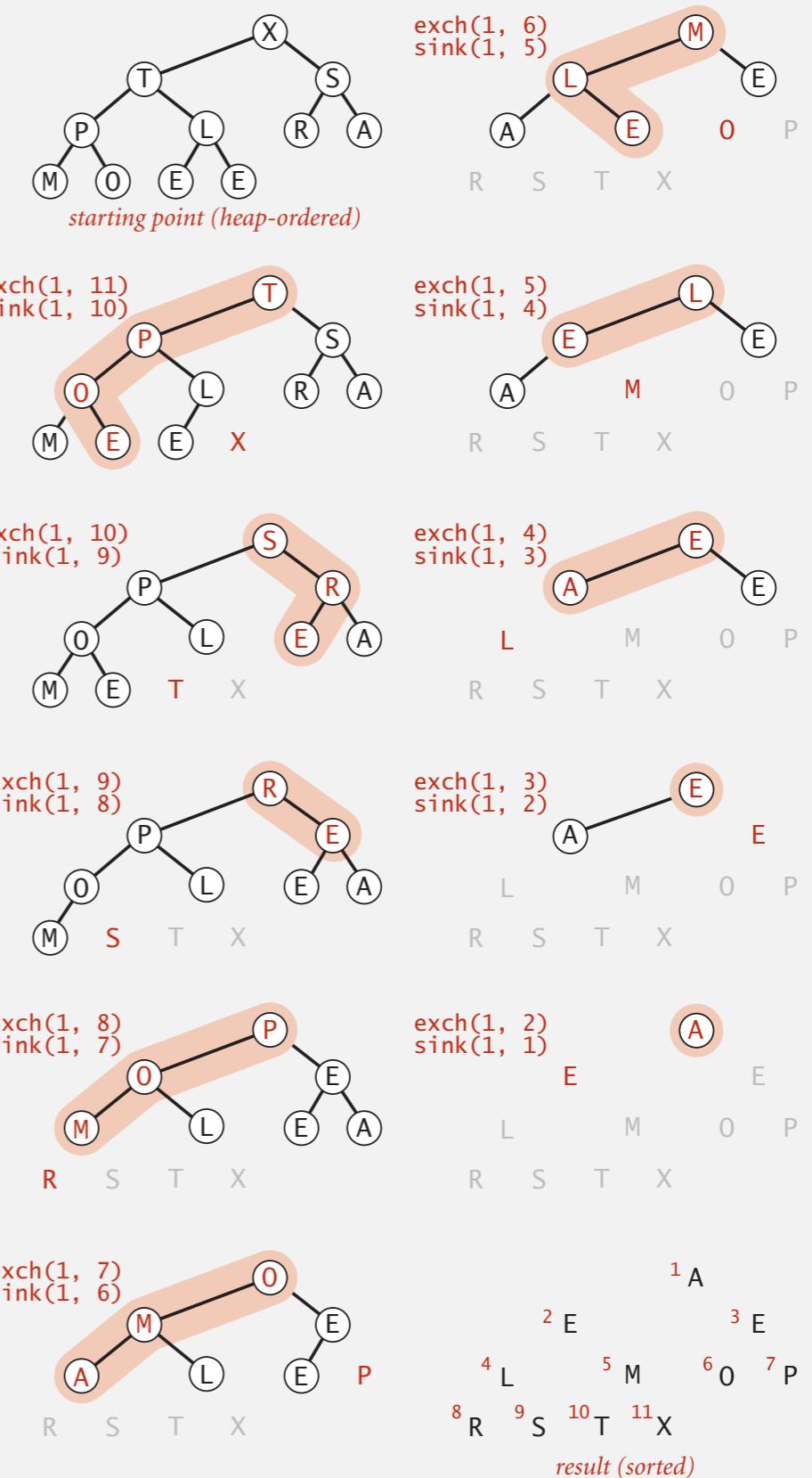


Heapsort: sortdown

Second pass.

- Remove the maximum, one at a time.
- Leave in array, instead of nulling out.

```
while (N > 1)
{
    swap(a, 1, N--);
    sink(a, 1, N);
}
```



Heapsort: Java implementation

```
public class Heap
{
    public static void sort(Comparable[] a)
    {
        int N = a.length;
        for (int k = N/2; k >= 1; k--)
            sink(a, k, N);
        while (N > 1)
        {
            swap(a, 1, N);
            sink(a, 1, --N);
        }
    }

    private static void sink(Comparable[] a, int k, int N)
    { /* as before */ }

    private static boolean less(Comparable[] a, int i, int j)
    { /* as before */ }
}

but make static (and pass arguments)
```

but convert from 1-based
indexing to 0-base indexing

```
private static void swap(Object[] a, int i, int j)
```

Heapsort: trace

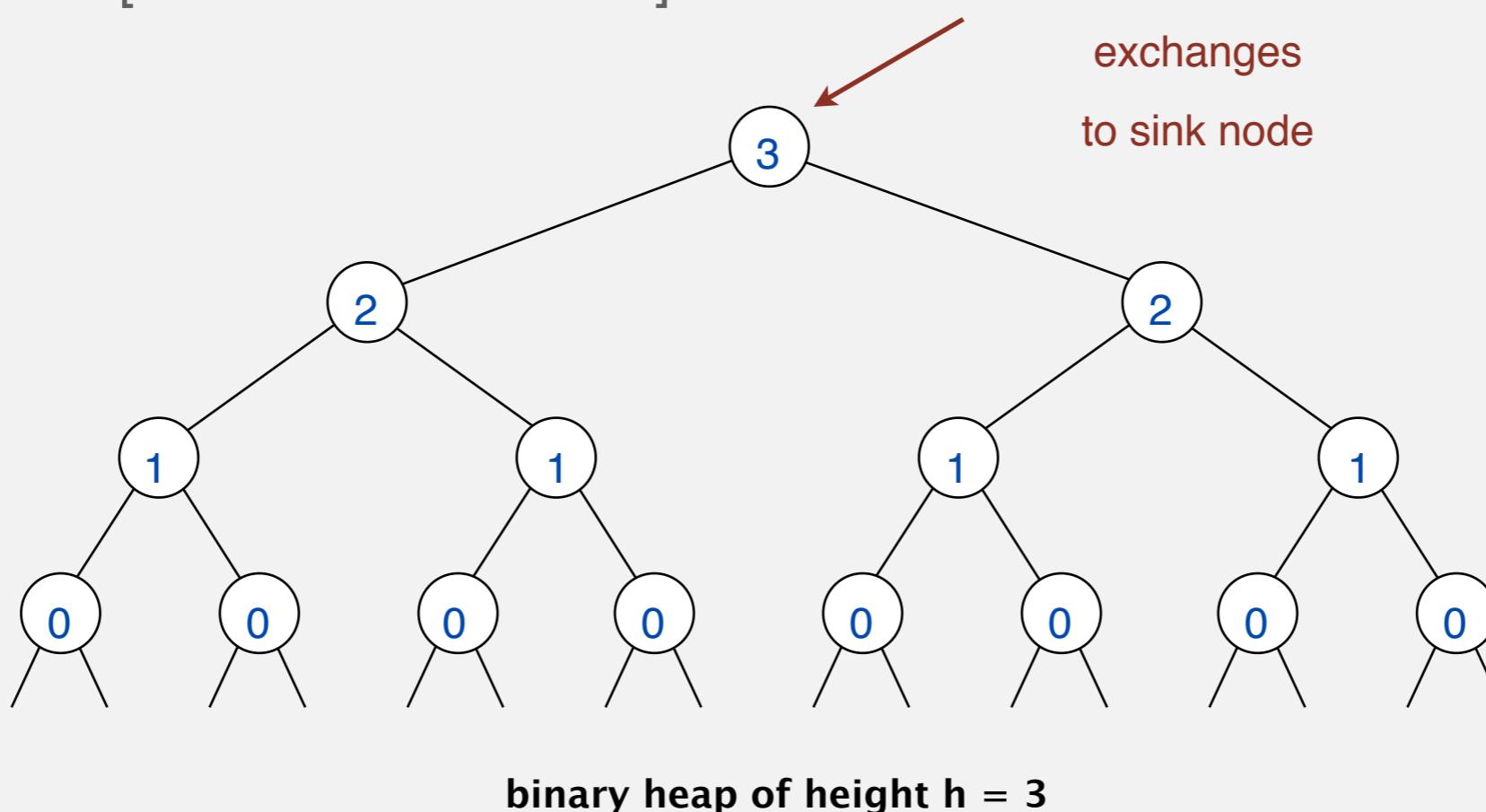
		a[i]											
N	k	0	1	2	3	4	5	6	7	8	9	10	11
<i>initial values</i>		S	0	R	T	E	X	A	M	P	L	E	
11	5	S	0	R	T	L	X	A	M	P	E	E	
11	4	S	0	R	T	L	X	A	M	P	E	E	
11	3	S	0	X	T	L	R	A	M	P	E	E	
11	2	S	T	X	P	L	R	A	M	O	E	E	
11	1	X	T	S	P	L	R	A	M	O	E	E	
<i>heap-ordered</i>		X	T	S	P	L	R	A	M	O	E	E	
10	1	T	P	S	O	L	R	A	M	E	E	X	
9	1	S	P	R	O	L	E	A	M	E	T	X	
8	1	R	P	E	O	L	E	A	M	S	T	X	
7	1	P	O	E	M	L	E	A	R	S	T	X	
6	1	O	M	E	A	L	E	P	R	S	T	X	
5	1	M	L	E	A	E	O	P	R	S	T	X	
4	1	L	E	E	A	M	O	P	R	S	T	X	
3	1	E	A	E	L	M	O	P	R	S	T	X	
2	1	E	A	E	L	M	O	P	R	S	T	X	
1	1	A	E	E	L	M	O	P	R	S	T	X	
<i>sorted result</i>		A	E	E	L	M	O	P	R	S	T	X	

Heapsort trace (array contents just after each sink)

Heapsort: mathematical analysis

Proposition. Heap construction uses $\leq 2N$ compares and $\leq N$ exchanges.

Pf sketch. [assume $N = 2^{h+1} - 1$]



$$\begin{aligned} h + 2(h-1) + 4(h-2) + 8(h-3) + \dots + 2^h(0) &\leq 2^{h+1} \\ &= N \end{aligned}$$

Heapsort: mathematical analysis

Proposition. Heap construction uses $\leq 2N$ compares and $\leq N$ exchanges.

Sortdown is basically the same as N calls of `deleteMax`. As we know, each of these is $O(\lg N)$.

Proposition. Heapsort uses $\leq 2N \lg N$ compares and exchanges.



algorithm can be improved to $\sim 1 N \lg N$

Significance. In-place sorting algorithm with $N \log N$ worst-case.

- Mergesort: no, linear extra space. ← in-place merge possible, not practical
- Quicksort: no, quadratic time in worst case. ← $N \log N$ worst-case quicksort possible, not practical
- Heapsort: yes!

Bottom line. Heapsort is optimal for both time and space, **but**:

- Inner loop longer than quicksort's.
- Makes poor use of cache.
- Not stable.



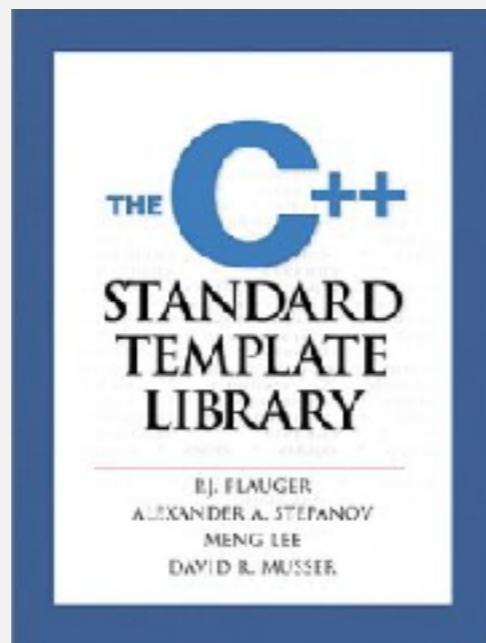
advanced tricks for improving

Introsort

Goal. As fast as quicksort in practice; $N \log N$ worst case, in place.

Introsort.

- Run quicksort.
- Cutoff to heapsort if stack depth exceeds $2 \lg N$.
- Cutoff to insertion sort for $N = 16$.



Introspective Sorting and Selection Algorithms

David R. Musser*
Computer Science Department
Rensselaer Polytechnic Institute, Troy, NY 12180
muss@cs.rpi.edu

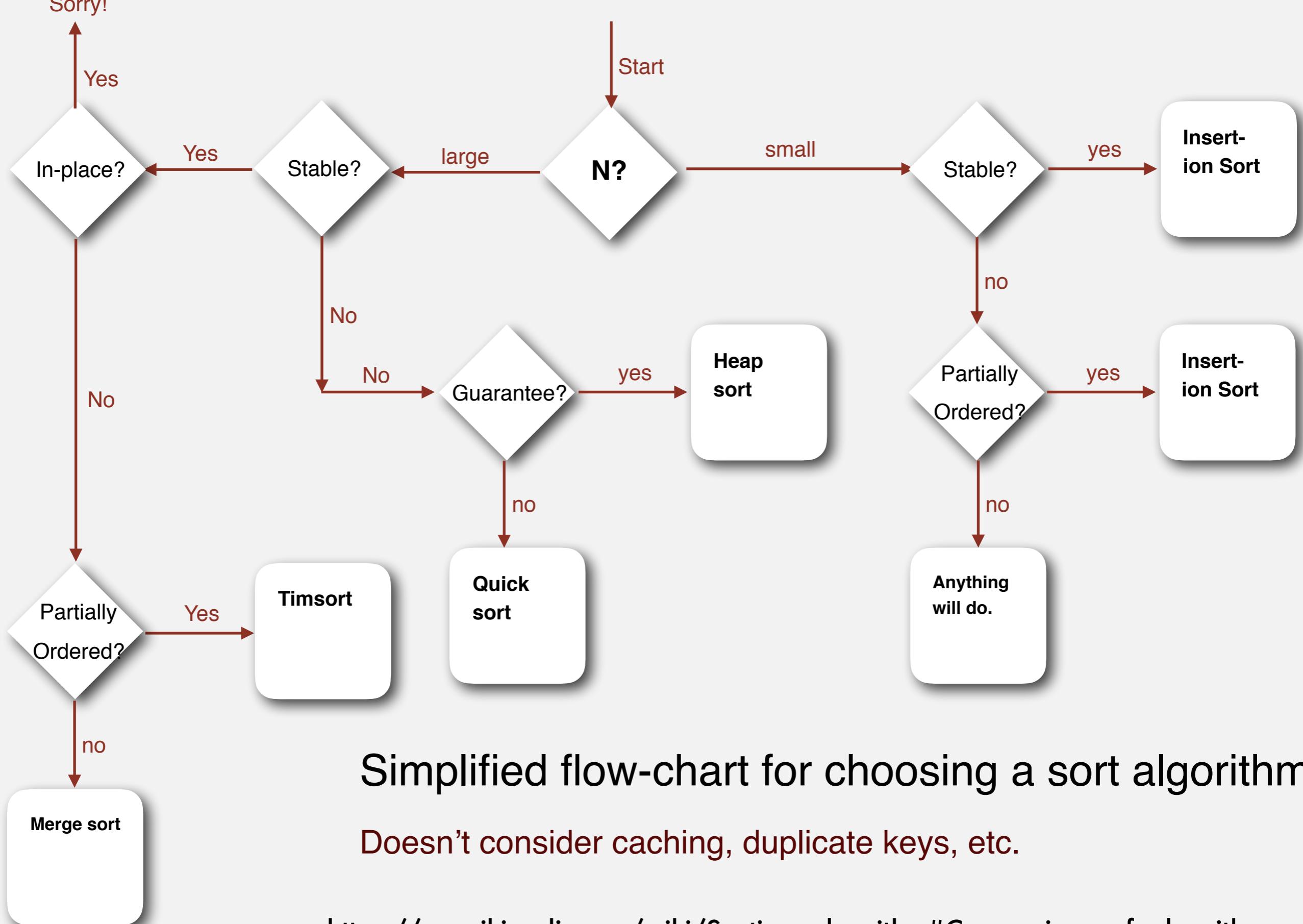
Abstract

Quicksort is the preferred in-place sorting algorithm in many contexts, since its average computing time on uniformly distributed inputs is $\Theta(N \log N)$ and it is in fact faster than most other sorting algorithms on most inputs. Its drawback is that its worst-case time bound is $\Theta(N^2)$. Previous attempts to protect against the worst case by improving the way quicksort chooses pivot elements for partitioning have increased the average computing time too much—one might as well use heapsort, which has a $\Theta(N \log N)$ worst-case time bound but is on the average 2 to 5 times slower than quicksort. A similar dilemma exists with selection algorithms (for finding the i -th largest element) based on partitioning. This paper describes a simple solution to this dilemma: limit the depth of partitioning, and for subproblems that exceed the limit switch to another algorithm with a better worst-case bound. Using heapsort as the “stopper” yields a sorting algorithm that is just as fast as quicksort in the average case but also has an $\Theta(N \log N)$ worst case time bound. For selection, a hybrid of Hoare’s FIND algorithm, which is linear on average but quadratic in the worst case, and the Blum-Floyd-Pratt-Rivest-Tarjan algorithm is as fast as Hoare’s algorithm in practice, yet has a linear worst-case time bound. Also discussed are issues of implementing the new algorithms as generic algorithms and accurately measuring their performance in the framework of the C++ Standard Template Library.

In the wild. C++ STL, Microsoft .NET Framework.

Sorting algorithms: summary

	In place?	Stable?	Best	Average	Worst	Remarks
selection	✓		$\frac{1}{2} N^2$	$\frac{1}{2} N^2$	$\frac{1}{2} N^2$	N exchanges
insertion	✓	✓	N	$\frac{1}{4} N^2$	$\frac{1}{2} N^2$	use for small N or partially ordered
shell	✓		$N \log_3 N$?	$c N^{3/2}$	tight code; subquadratic
merge		✓	$\frac{1}{2} N \lg N$	$N \lg N$	$N \lg N$	$N \log N$ guarantee; stable
timsort		✓	N	$N \lg N$	$N \lg N$	improves mergesort when preexisting order
quick	✓		$N \lg N$	$2 N \ln N$	$\frac{1}{2} N^2$	$N \log N$ probabilistic guarantee; fastest in practice
3-way quick	✓		N	$2 N \ln N$	$\frac{1}{2} N^2$	improves quicksort when duplicate keys
heap	✓		N	$2 N \lg N$	$2 N \lg N$	$N \log N$ guarantee; in-place
?	✓	✓	N	$N \lg N$	$N \lg N$	holy sorting grail



Algorithms

ROBERT SEDGEWICK | KEVIN WAYNE

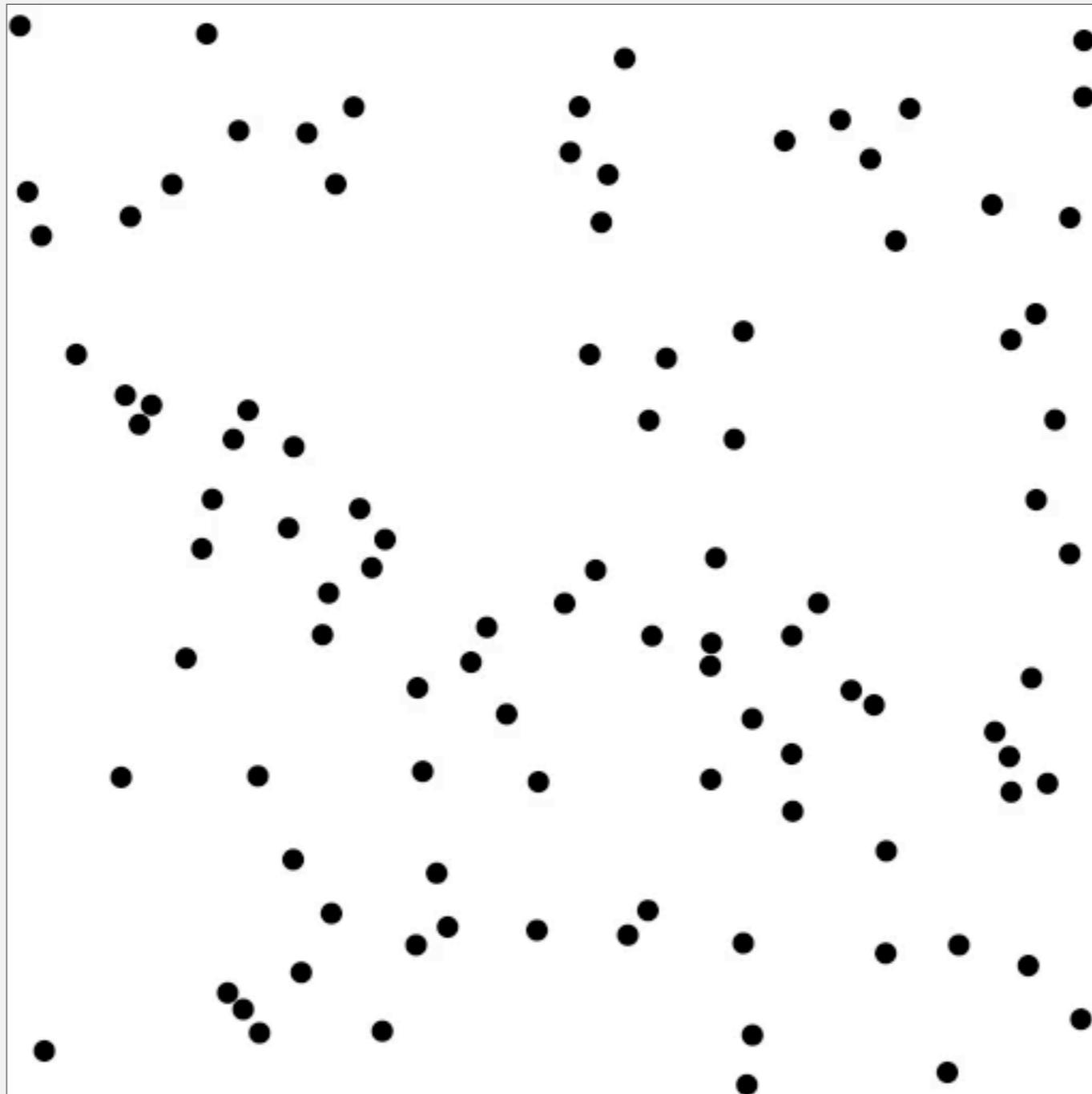
<http://algs4.cs.princeton.edu>

2.4 PRIORITY QUEUES

- ▶ *API and elementary implementations*
- ▶ *binary heaps*
- ▶ *heapsort*
- ▶ ***event-driven simulation***

Molecular dynamics simulation of hard discs

Goal. Simulate the motion of N moving particles that behave according to the laws of elastic collision.

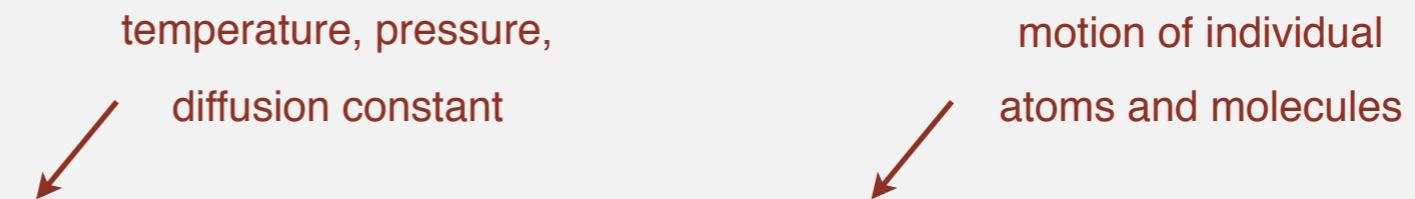


Molecular dynamics simulation of hard discs

Goal. Simulate the motion of N moving particles that behave according to the laws of elastic collision.

Hard disc model.

- Moving particles interact via elastic collisions with each other and walls.
- Each particle is a disc with known position, velocity, mass, and radius.
- No other forces.



Significance. Relates macroscopic observables to microscopic dynamics.

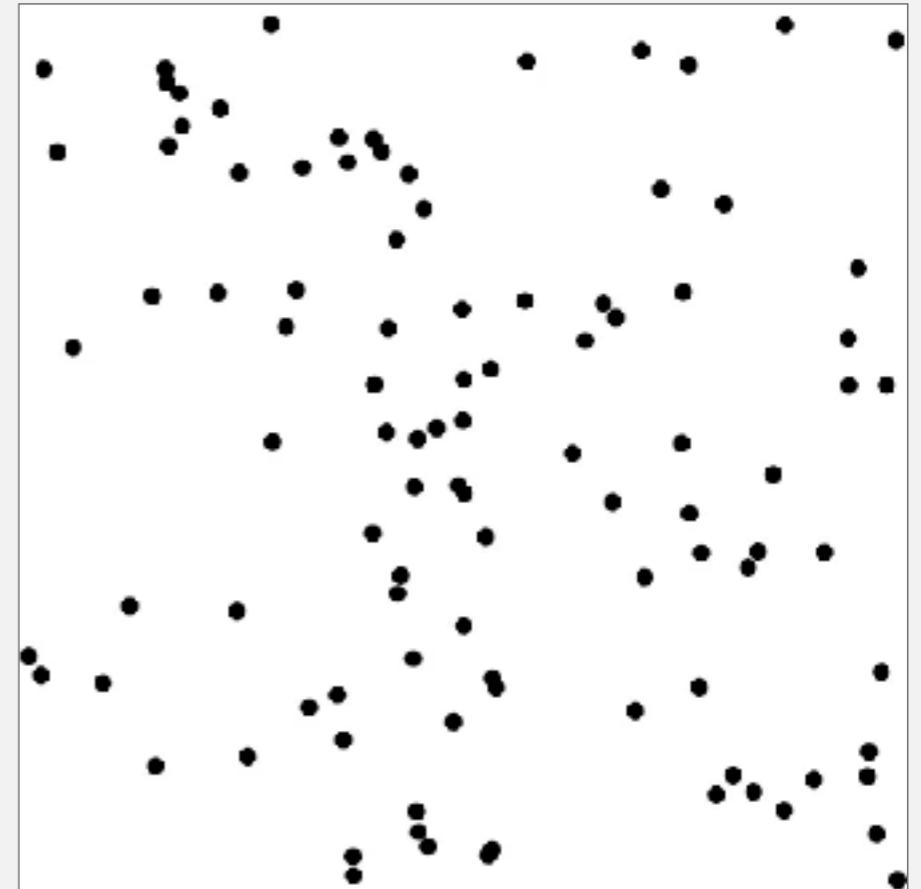
- Maxwell-Boltzmann: distribution of speeds as a function of temperature.
- Einstein: explain Brownian motion of pollen grains.

Warmup: bouncing balls

Time-driven simulation. N bouncing balls in the unit square.

```
public class BouncingBalls
{
    public static void main(String[] args)
    {
        int N = Integer.parseInt(args[0]);
        Ball[] balls = new Ball[N];
        for (int i = 0; i < N; i++)
            balls[i] = new Ball();
        while(true)
        {
            StdDraw.clear();
            for (int i = 0; i < N; i++)
            {
                balls[i].move(0.5);
                balls[i].draw();
            }
            StdDraw.show(50);
        }
    }
}
```

% java BouncingBalls 100



main simulation loop



Warmup: bouncing balls

```
public class Ball
{
    private double rx, ry;      // position
    private double vx, vy;      // velocity
    private final double radius; // radius
    public Ball(...)
    { /* initialize position and velocity */ }

    public void move(double dt)
    {
        if ((rx + vx*dt < radius) || (rx + vx*dt > 1.0 - radius)) { vx = -vx; }
        if ((ry + vy*dt < radius) || (ry + vy*dt > 1.0 - radius)) { vy = -vy; }
        rx = rx + vx*dt;
        ry = ry + vy*dt;
    }

    public void draw()
    { StdDraw.filledCircle(rx, ry, radius); }
}
```

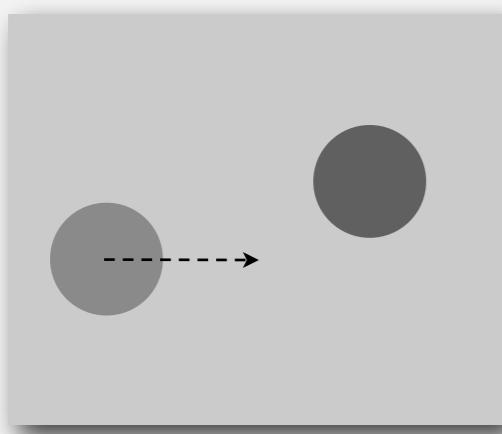
check for collision with walls



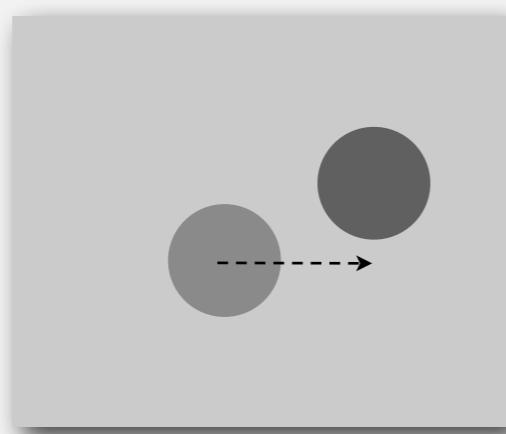
- CS problems: which object does the check? too many checks?

Time-driven simulation

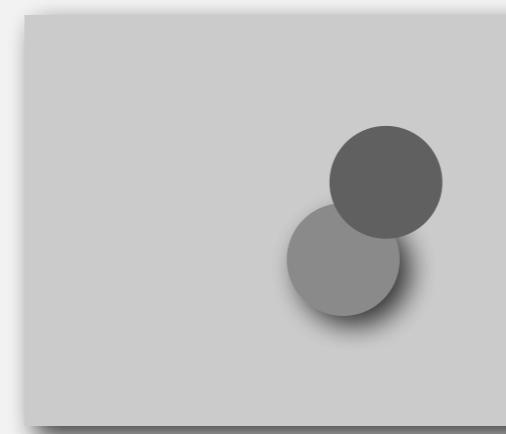
- Discretize time in quanta of size dt .
- Update the position of each particle after every dt units of time, and check for overlaps.
- If overlap, roll back the clock to the time of the collision, update the velocities of the colliding particles, and continue the simulation.



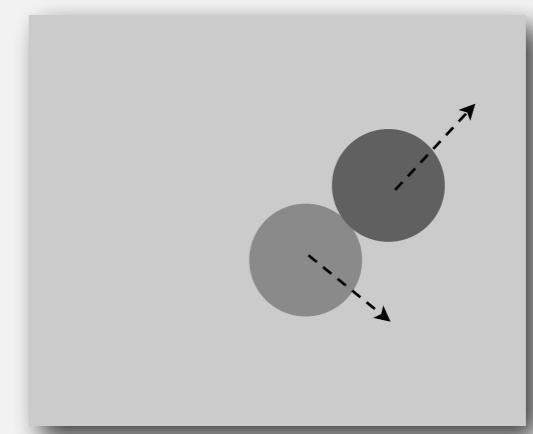
t



$t + dt$



$t + 2 dt$
(collision detected)

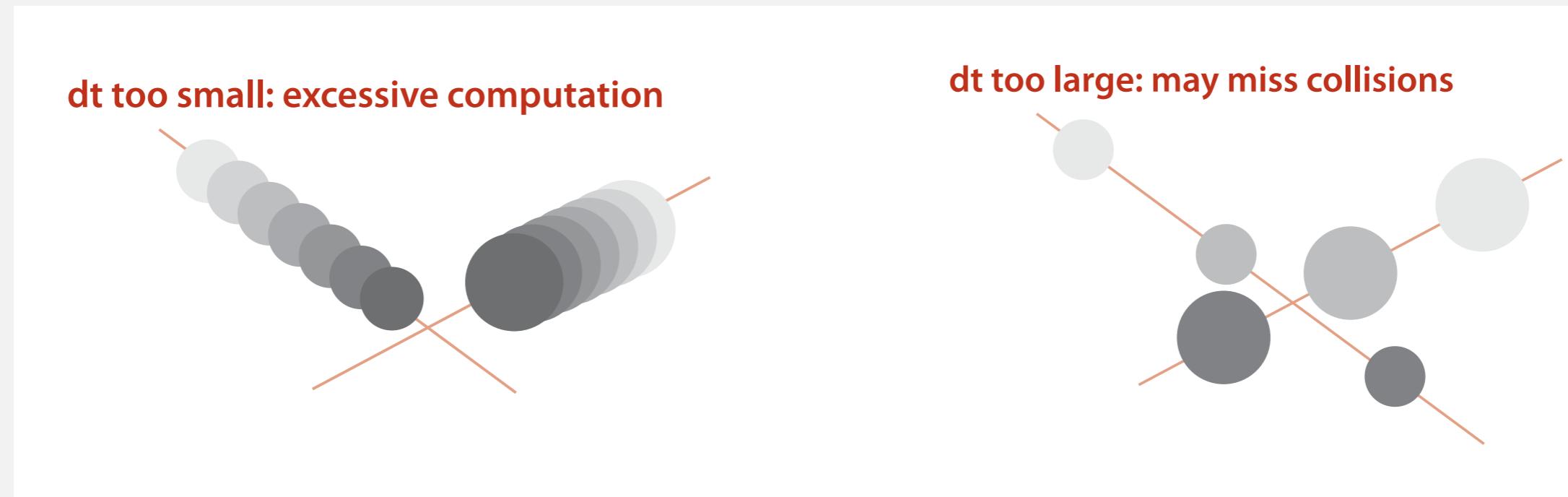


$t + \Delta t$
(roll back clock)

Time-driven simulation

Main drawbacks.

- $\sim N^2 / 2$ overlap checks per time quantum.
- Simulation is too slow if dt is very small.
- May miss collisions if dt is too large.
(if colliding particles fail to overlap when we are looking)



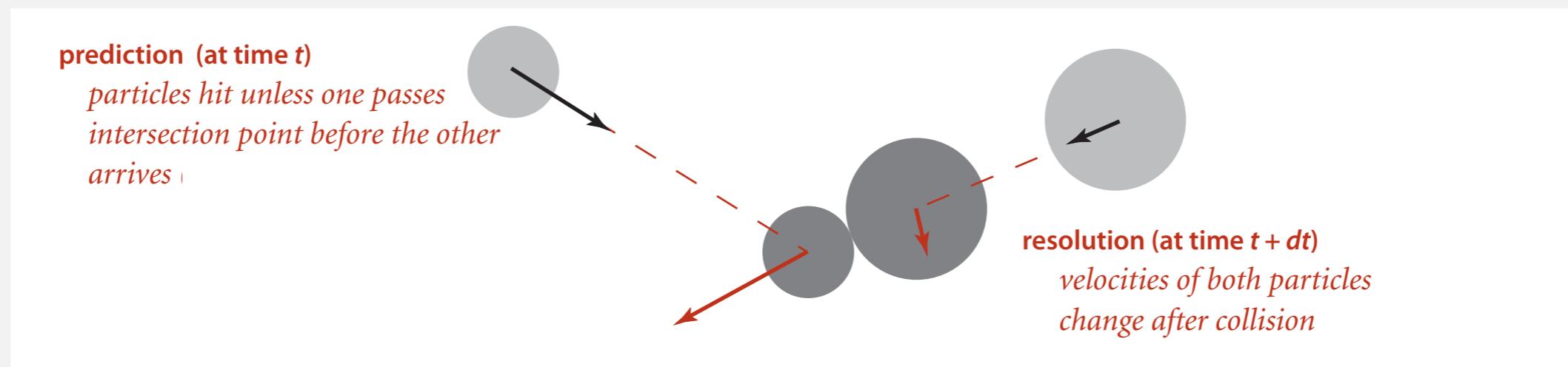
Event-driven simulation

Change state only when something happens.

- Between collisions, particles move in straight-line trajectories.
- Focus only on times when collisions occur.
- Maintain PQ of collision events, prioritized by time.
- Remove the min = get next collision.

Collision prediction. Given position, velocity, and radius of a particle, when will it collide next with a wall or another particle?

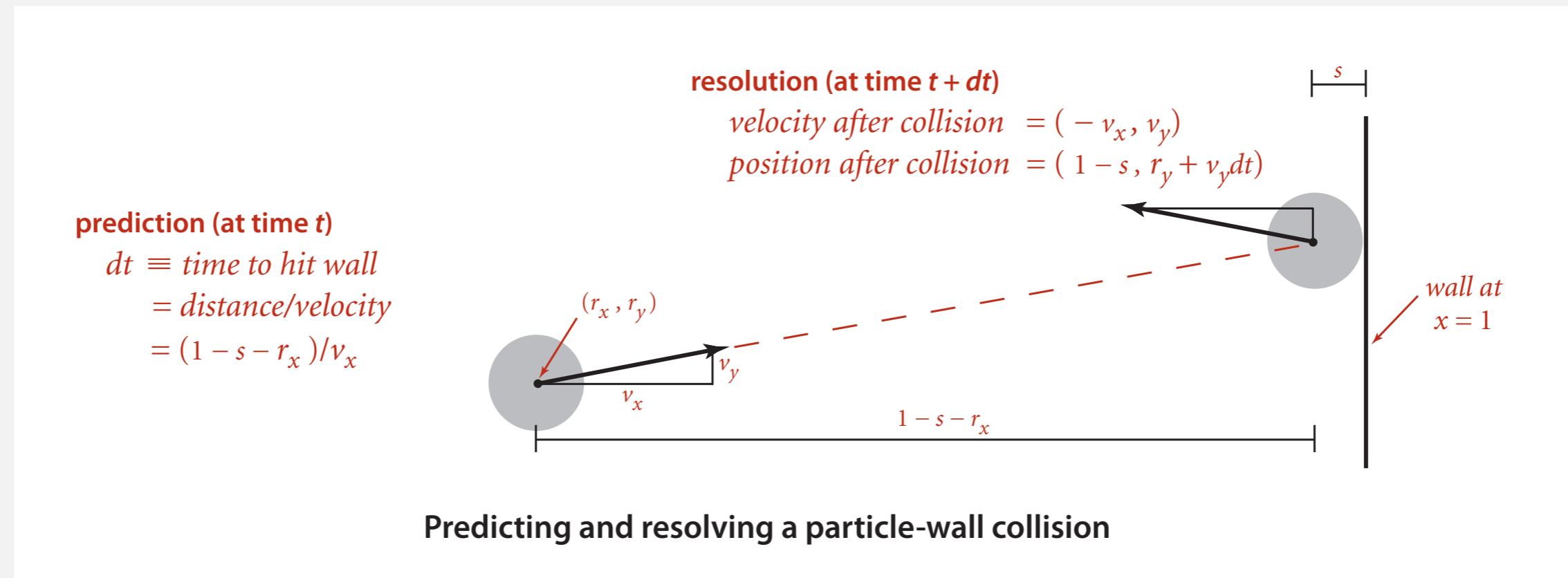
Collision resolution. If collision occurs, update colliding particle(s) according to laws of elastic collisions.



Particle-wall collision

Collision prediction and resolution.

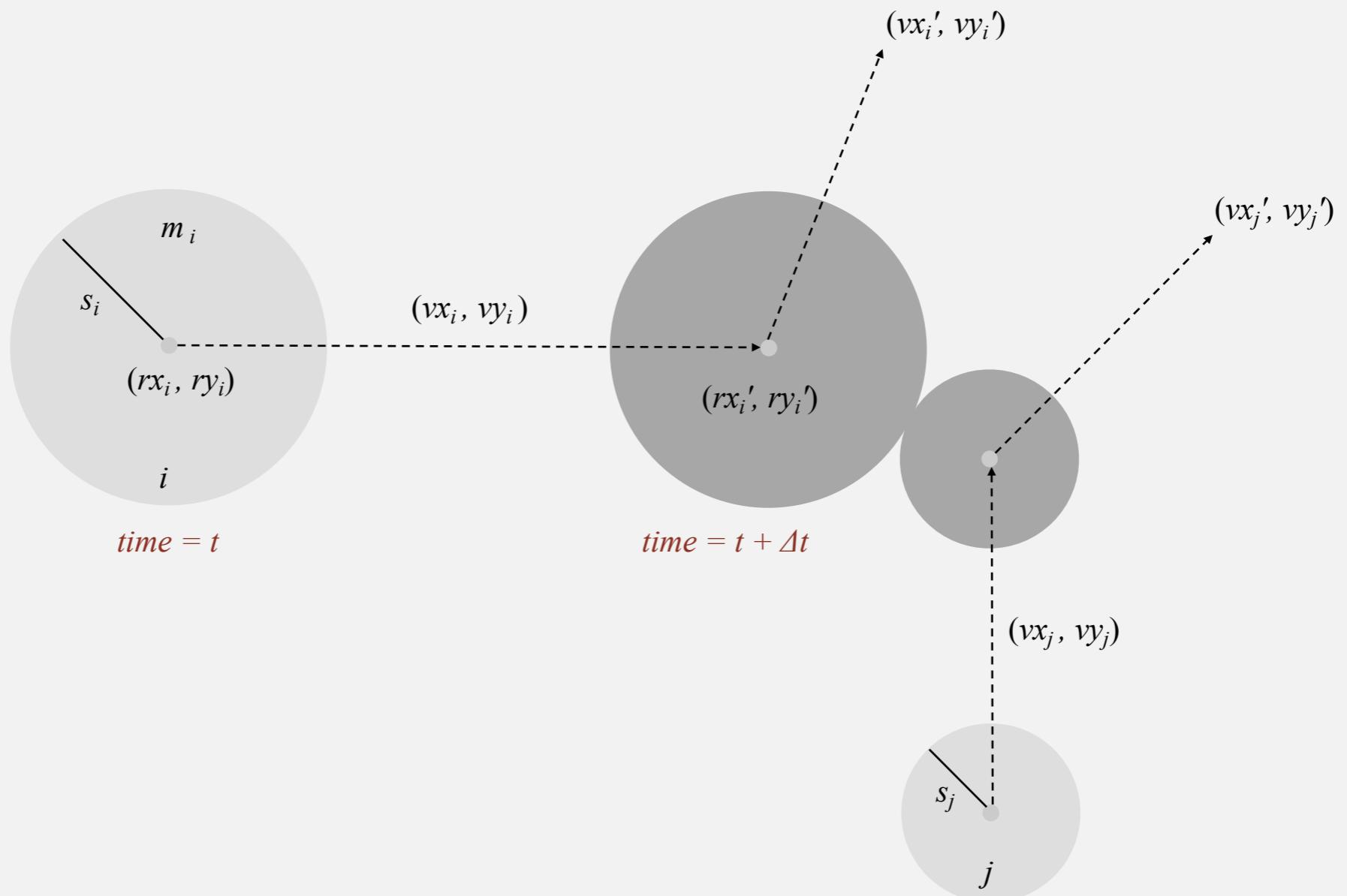
- Particle of radius s at position (r_x, r_y) .
- Particle moving in unit box with velocity (v_x, v_y) .
- Will it collide with a vertical wall? If so, when?



Particle-particle collision prediction

Collision prediction.

- Particle i : radius s_i , position (rx_i, ry_i) , velocity (vx_i, vy_i) .
- Particle j : radius s_j , position (rx_j, ry_j) , velocity (vx_j, vy_j) .
- Will particles i and j collide? If so, when?



Particle-particle collision prediction

Collision prediction.

- Particle i : radius s_i , position (rx_i, ry_i) , velocity (vx_i, vy_i) .
- Particle j : radius s_j , position (rx_j, ry_j) , velocity (vx_j, vy_j) .
- Will particles i and j collide? If so, when?

$$\Delta t = \begin{cases} \infty & \text{if } \Delta v \cdot \Delta r \geq 0 \\ \infty & \text{if } d < 0 \\ -\frac{\Delta v \cdot \Delta r + \sqrt{d}}{\Delta v \cdot \Delta v} & \text{otherwise} \end{cases}$$

$$d = (\Delta v \cdot \Delta r)^2 - (\Delta v \cdot \Delta v) (\Delta r \cdot \Delta r - \sigma^2) \quad \sigma = \sigma_i + \sigma_j$$

$$\Delta v = (\Delta vx, \Delta vy) = (vx_i - vx_j, vy_i - vy_j)$$

$$\Delta r = (\Delta rx, \Delta ry) = (rx_i - rx_j, ry_i - ry_j)$$

$$\Delta v \cdot \Delta v = (\Delta vx)^2 + (\Delta vy)^2$$

$$\Delta r \cdot \Delta r = (\Delta rx)^2 + (\Delta ry)^2$$

$$\Delta v \cdot \Delta r = (\Delta vx)(\Delta rx) + (\Delta vy)(\Delta ry)$$

Important note: This is physics, so we won't be testing you on it!

Particle-particle collision resolution

Collision resolution. When two particles collide, how does velocity change?

$$\begin{aligned} vx_i' &= vx_i + Jx / m_i \\ vy_i' &= vy_i + Jy / m_i \\ vx_j' &= vx_j - Jx / m_j \\ vy_j' &= vy_j - Jy / m_j \end{aligned}$$

Newton's second law
(momentum form)

$$J_x = \frac{J \Delta r_x}{\sigma}, \quad J_y = \frac{J \Delta r_y}{\sigma}, \quad J = \frac{2 m_i m_j (\Delta v \cdot \Delta r)}{\sigma(m_i + m_j)}$$

impulse due to normal force

(conservation of energy, conservation of momentum)

Important note: This is physics, so we won't be testing you on it!

Particle data type skeleton

```
public class Particle
{
    private double rx, ry;      // position
    private double vx, vy;      // velocity
    private final double radius; // radius
    private final double mass;   // mass
    private int count;          // number of collisions
```

```
public Particle(...) { }
```

```
public void move(double dt) { }
```

```
public void draw() { }
```

```
public double timeToHit(Particle that) { }
```

```
public double timeToHitVerticalWall() { }
```

```
public double timeToHitHorizontalWall() { }
```

```
public void bounceOff(Particle that) { }
```

```
public void bounceOffVerticalWall() { }
```

```
public void bounceOffHorizontalWall() { }
```

predict collision with
particle or wall

resolve collision with
particle or wall

Particle-particle collision and resolution implementation

```
public double timeToHit(Particle that)
{
    if (this == that) return INFINITY;
    double dx = that.rx - this.rx, dy = that.ry - this.ry;
    double dvx = that.vx - this.vx; dvy = that.vy - this.vy;
    double dvdr = dx*dvx + dy*dvy;
    if( dvdr > 0) return INFINITY;
    double dvdv = dvx*dvx + dvy*dvy;
    double drdr = dx*dx + dy*dy;
    double sigma = this.radius + that.radius;
    double d = (dvdr*dvdr) - dvdv * (drdr - sigma*sigma);
    if (d < 0) return INFINITY;
    return -(dvdr + Math.sqrt(d)) / dvdv;
}
```

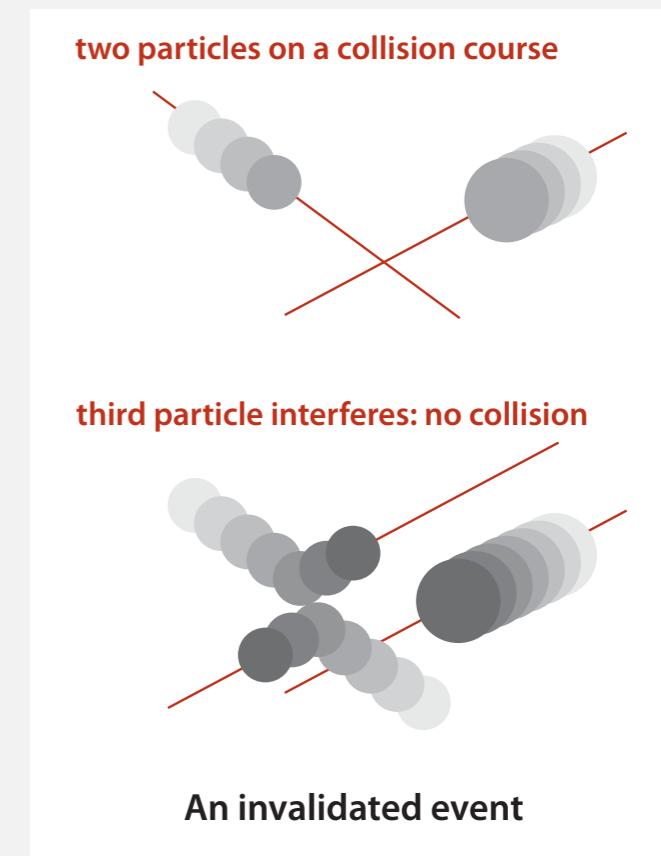
```
public void bounceOff(Particle that)
{
    double dx = that.rx - this.rx, dy = that.ry - this.ry;
    double dvx = that.vx - this.vx, dvy = that.vy - this.vy;
    double dvdr = dx*dvx + dy*dvy;
    double dist = this.radius + that.radius;
    double J = 2 * this.mass * that.mass * dvdr / ((this.mass + that.mass) * dist);
    double Jx = J * dx / dist;
    double Jy = J * dy / dist;
    this.vx += Jx / this.mass;
    this.vy += Jy / this.mass; Important note: This is physics, so we won't be testing you on it!
    that.vx -= Jx / that.mass;
    that.vy -= Jy / that.mass;
}
```

Collision system: event-driven simulation main loop

Initialization.

- Fill PQ with all potential particle-wall collisions.
- Fill PQ with all potential particle-particle collisions.

“potential” since collision may not happen if
some other collision intervenes



Main loop.

- Delete the impending event from PQ (min priority = t).
- If the event has been invalidated, ignore it.
- Advance all particles to time t , on a straight-line trajectory.
- Update the velocities of the colliding particle(s).
- Predict future particle-wall and particle-particle collisions involving the colliding particle(s) and insert events onto PQ.

Event data type

Conventions.

- Neither particle null \Rightarrow particle-particle collision.
- One particle null \Rightarrow particle-wall collision.
- Both particles null \Rightarrow redraw event.

```
private class Event implements Comparable<Event>
{
    private double time;      // time of event
    private Particle a, b;    // particles involved in event
    private int countA, countB; // collision counts for a and b
```

```
public Event(double t, Particle a, Particle b) { }
```

```
public int compareTo(Event that)
{ return this.time - that.time; }
```

```
public boolean isValid()
{ }
```

← create event
← ordered by time
← invalid if
intervening collision

Collision system implementation: skeleton

```
public class CollisionSystem
{
    private MinPQ<Event> pq;      // the priority queue
    private double t = 0.0;         // simulation clock time
    private Particle[] particles; // the array of particles

    public CollisionSystem(Particle[] particles) { }

    private void predict(Particle a)
    {
        if (a == null) return;
        for (int i = 0; i < N; i++)
        {
            double dt = a.timeToHit(particles[i]);
            pq.insert(new Event(t + dt, a, particles[i]));
        }
        pq.insert(new Event(t + a.timeToHitVerticalWall() , a, null));
        pq.insert(new Event(t + a.timeToHitHorizontalWall(), null, a));
    }

    private void redraw() {}
```

add to PQ all particle-wall and particle-particle collisions involving this particle

Collision system implementation: main event-driven simulation loop

```
public void simulate()
{
    pq = new MinPQ<Event>();
    for(int i = 0; i < N; i++) predict(particles[i]);
    pq.insert(new Event(0, null, null));
```

initialize PQ with collision events and redraw event

```
while(!pq.isEmpty())
{
    Event event = pq.delMin();
    if(!event.isValid()) continue;
    Particle a = event.a;
    Particle b = event.b;
```

get next event

```
    for(int i = 0; i < N; i++)
        particles[i].move(event.time - t);
    t = event.time;
```

update positions and time

```
    if      (a != null && b != null) a.bounceOff(b);
    else if (a != null && b == null) a.bounceOffVerticalWall()
    else if (a == null && b != null) b.bounceOffHorizontalWall();
    else if (a == null && b == null) redraw();
```

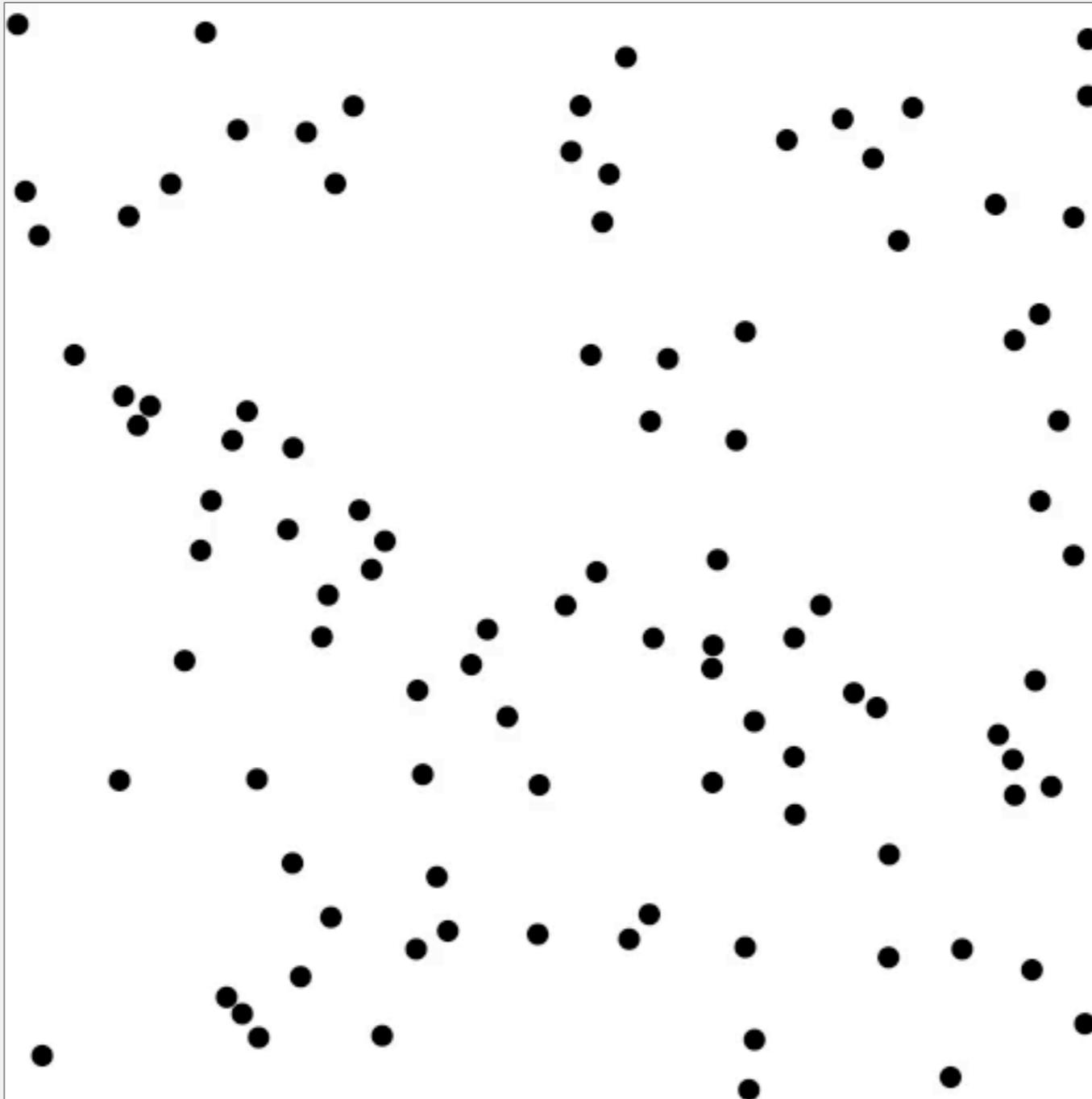
process event

```
    predict(a);
    predict(b);
```

predict new events based on changes

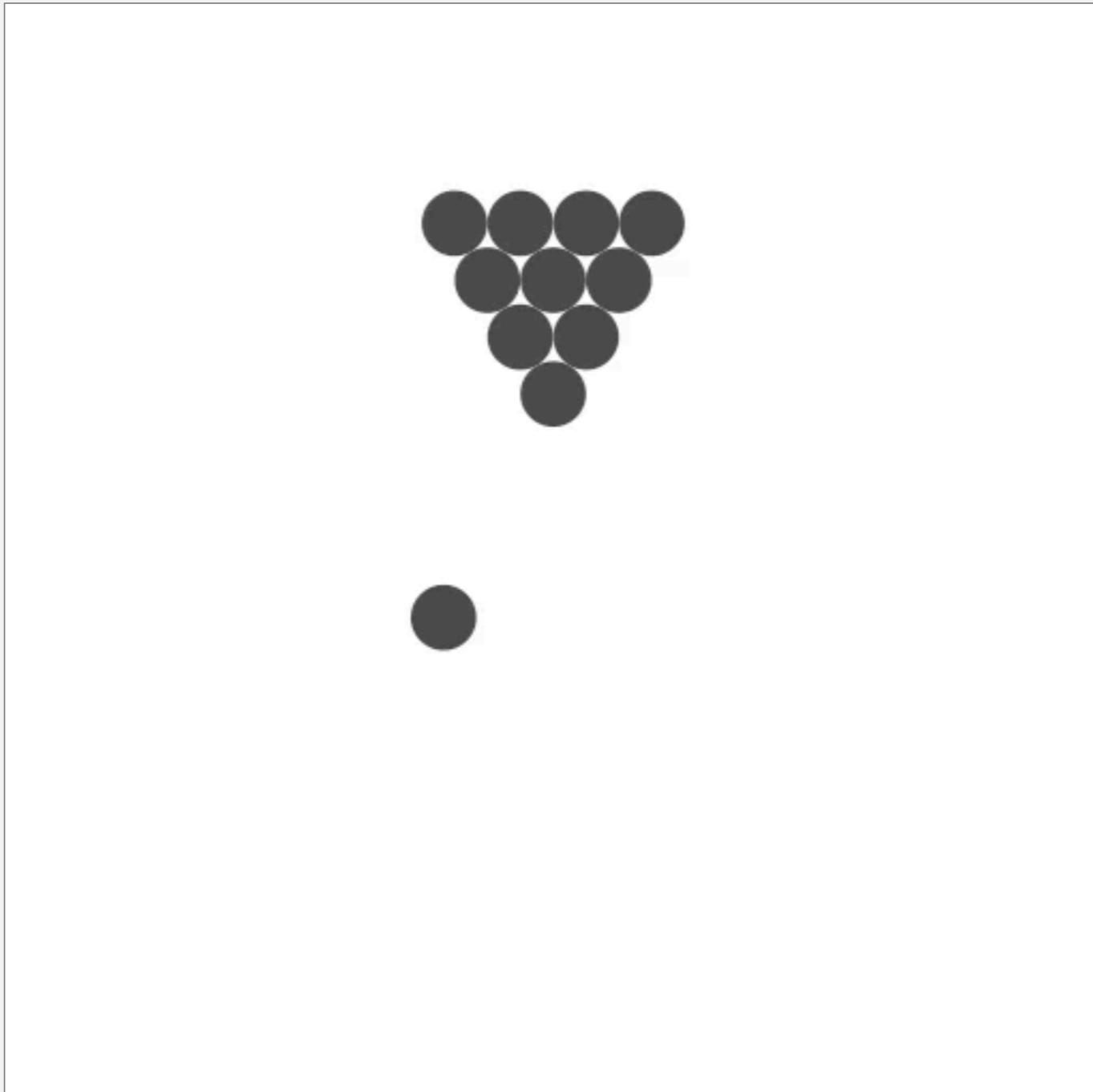
Particle collision simulation example 1

% java CollisionSystem 100



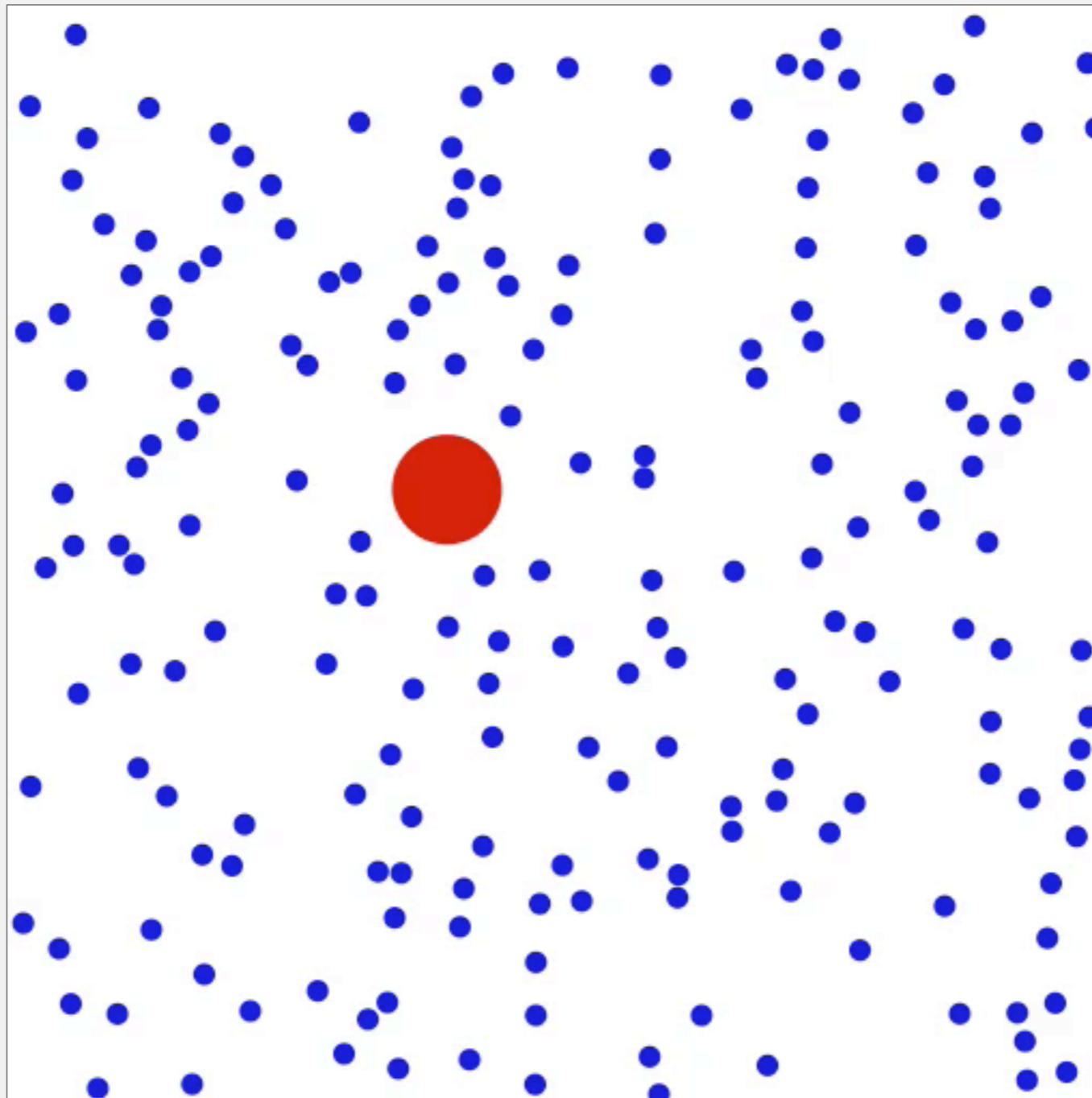
Particle collision simulation example 2

```
% java CollisionSystem < billiards.txt
```



Particle collision simulation example 3

```
% java CollisionSystem < brownian.txt
```



Particle collision simulation example 4

% java CollisionSystem < diffusion.txt

