Advanced Algorithmics (6EAP)

MTAT.03.238

Linear structures, sorting, searching, etc

Jaak Vilo

2012 Spring
Linear, sequential, ordered, list ...
Physical ordered list ~ array

• Memory /address/
  – Garbage collection

• Files (character/byte list/lines in text file,...)

• Disk
  – Disk fragmentation
Lists: Array

L = int[MAX_SIZE]

L[2] = 7
Lists: Array

$L = \text{int}[\text{MAX}\_\text{SIZE}]$

$L[2]=7$

$L[3]=7$
Multiple lists, 2-D-arrays, etc...
Linear Lists

• Operations which one may want to perform on a linear list of $n$ elements include:
  
  – gain access to the $k$th element of the list to examine and/or change the contents
  – insert a new element before or after the $k$th element
  – delete the $k$th element of the list

Abstract Data Type (ADT)

• High-level definition of data types

• An ADT specifies
  – A collection of data
  – A set of operations on the data or subsets of the data

• ADT does not specify how the operations should be implemented

• Examples
  – vector, list, stack, queue, deque, priority queue, table (map), associative array, set, graph, digraph
ADT

• A **datatype** is a **set of values** and an associated **set of operations**
• A datatype is **abstract** iff it is completely described by its set of operations regardless of its implementation
• This means that it is possible to change the implementation of the datatype without changing its use
• The datatype is thus described by a set of procedures
• These operations are the only thing that a user of the abstraction can assume
Abstract data types:

• Dictionary (key,value)
• Stack (LIFO)
• Queue (FIFO)
• Queue (double-ended)
• Priority queue (fetch highest-value object)
• ...

DicHonary

Stack (LIFO)

Queue (FIFO)

Queue (double-ended)

Priority queue (fetch highest-value object)
Dictionary

• **Container of key-element (k,e) pairs**
  - Required operations:
    - `insert( k,e )`,
    - `remove( k )`,
    - `find( k )`,
    - `isEmpty()`
  - May also support (when an order is provided):
    - `closestKeyBefore( k )`,
    - `closestElemAfter( k )`
• **Note:** No duplicate keys
Abstract data types

- Container
- Deque
- Map/Associative array/Dictionary
- Multimap
- Multiset
- Priority queue
- Queue
- Set
- Stack
- String
- Tree
- Graph
- Hash
Some data structures for Dictionary ADT

• Unordered
  – Array
  – Sequence/list

• Ordered
  – Array
  – Sequence (Skip Lists)
  – Binary Search Tree (BST)
  – AVL trees, red-black trees
  – (2; 4) Trees
  – B-Trees

• Valued
  – Hash Tables
  – Extendible Hashing
Primitive & composite types

**Primitive types**
- **Boolean** (for boolean values True/False)
- **Char** (for character values)
- **int** (for integral or fixed-precision values)
- **Float** (for storing real number values)
- **Double** (a larger size of type float)
- **String** (for string of chars)
- **Enumerated** type

**Composite types**
- **Array**
- **Record** (also called tuple or struct)
- **Union**
- Tagged union (also called a variant, variant record, discriminated union, or disjoint union)
- Plain old data structure
# Linear data structures

## Arrays
- Array
- Bidirectional map
- Bit array
- Bit field
- Bitboard
- Bitmap
- Circular buffer
- Control table
- Image
- Dynamic array
  - Gap buffer
  - Hashed array tree
  - Heightmap
  - Lookup table
  - Matrix
  - Parallel array
  - Sorted array
  - Sparse array
  - Sparse matrix
  - Iliffe vector
  - Variable-length array

## Lists
- Doubly linked list
- Linked list
- Self-organizing list
- Skip list
- Unrolled linked list
- VList
- Xor linked list
- Zipper
- Doubly connected edge list
Trees

**Binary trees**
- AA tree
- AVL tree
- Binary search tree
- Binary tree
- Cartesian tree
- Pagoda
- Randomized binary search tree
- Red-black tree
- Rope
- Scapegoat tree
- Self-balancing binary search tree
- Splay tree
- T-tree
- Tango tree
- Threaded binary tree
- Top tree

**Heaps**
- Heap
- Binary heap
- Binomial heap
- Fibonacci heap
- AF-heap

**B-trees**
- B-tree
- B+ tree
- B*-tree
- B sharp tree
- Dancing tree
- 2-3 tree
- 2-3-4 tree
- Queap
- Fusion tree
- Bx-tree

**Tries**
- Trie
- Radix tree
- Suffix tree
- Suffix array
- Compressed suffix array
- FM-index
- Generalised suffix tree
- B-trie
- Judy array

**Multiway trees**
- 2-3 heap
- Soft heap
- Pairing heap
- Leftist heap
- Treap
- Beap
- Skew heap
- Ternary heap
- D-ary heap

**Space-partitioning trees**
- X-fast trie
- Y-fast trie
- C-trie

**Application-specific trees**
- Treap
- Weight-balanced tree

- Space-partitioning trees
  - Segment tree
  - Interval tree
  - Range tree
  - Bin
  - Kd-tree
  - Implicit kd-tree
  - Min/max kd-tree
  - Adaptive k-d tree
  - Kdb tree
  - Quadtree
  - Octree
  - Linear octree
  - Z-order
  - UB-tree
  - R-tree
  - R+ tree
  - R* tree
  - Hilbert R-tree
  - X-tree
  - Metric tree
  - Cover tree
  - M-tree
  - VP-tree
  - BK-tree
  - Bounding interval hierarchy
  - BSP tree
  - Rapidly-exploring random tree
  - Syntax tree
  - Abstract syntax tree
  - Parse tree
  - Decision tree
  - Alternating decision tree
  - Minimax tree
  - Expectiminimax tree
  - Finger tree
Hashes, Graphs, Other

**Hashes**
- Bloom filter
- Distributed hash table
- Hash array mapped trie
- Hash list
- Hash table
- Hash tree
- Hash trie
- Koorde
- Prefix hash tree

**Graphs**
- Adjacency list
- Adjacency matrix
- Graph-structured stack
- Scene graph
- Binary decision diagram
- Zero suppressed decision diagram
- And-inverter graph
- Directed graph

**Directed acyclic graph**
**Propositional directed acyclic graph**
**Multigraph**
**Hypergraph**

**Other**
- Lightmap
- Winged edge
- Quad-edge
- Routing table
- Symbol table
Lists: Array

size = MAX_SIZE - 1

Insert 8 after L[2]

Delete last
Lists: Array

- Insert: $O(n)$
- Delete: $O(n)$
- Access i: $O(1)$
- Insert to end: $O(1)$
- Delete from end: $O(1)$
- Search: $O(n)$

- Insert 8 after L[2]
- Delete last
Linear Lists

• Other operations on a linear list may include:
  – determine the number of elements
  – search the list
  – sort a list
  – combine two or more linear lists
  – split a linear list into two or more lists
  – make a copy of a list
Stack

• push(x) -- add to end (add to top)
• pop() -- fetch from end (top)

• O(1) in all reasonable cases 😊

• LIFO – Last In, First Out
Linked lists

Singly linked

Doubly linked
Linked lists: add
Linked lists: delete
(+ garbage collection?)

head

tail

size
# Operations

- **Array indexed from 0 to \( n - 1 \):**

<table>
<thead>
<tr>
<th>Operation</th>
<th>( k = 1 )</th>
<th>( 1 &lt; k &lt; n )</th>
<th>( k = n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>access/change the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(1) )</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>insert before or after the ( k )th element</td>
<td>( O(n) )</td>
<td>( O(n) )</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>delete the ( k )th element</td>
<td>( O(n) )</td>
<td>( O(n) )</td>
<td>( O(1) )</td>
</tr>
</tbody>
</table>

- **Singly-linked list with head and tail pointers**

<table>
<thead>
<tr>
<th>Operation</th>
<th>( k = 1 )</th>
<th>( 1 &lt; k &lt; n )</th>
<th>( k = n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>access/change the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(n) )</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>insert before or after the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(n) ) ( O(1)^1 )</td>
<td>( O(n) ) ( O(1) )</td>
</tr>
<tr>
<td>delete the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(n) )</td>
<td>( O(n) )</td>
</tr>
</tbody>
</table>

1 under the assumption we have a pointer to the \( k \)th node, \( O(n) \) otherwise
Improving Run-Time Efficiency

• We can improve the run-time efficiency of a linked list by using a doubly-linked list:

  Singly-linked list:
  
  ![Singly-linked list diagram]

  Doubly-linked list:
  
  ![Doubly-linked list diagram]

  – Improvements at operations requiring access to the previous node
  – Increases memory requirements...
# Improving Efficiency

## Singly-linked list:

<table>
<thead>
<tr>
<th>Action</th>
<th>$k = 1$</th>
<th>$1 &lt; k &lt; n$</th>
<th>$k = n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>access/change the $k$th element</td>
<td>$O(1)$</td>
<td>$O(n)$</td>
<td>$O(1)$</td>
</tr>
<tr>
<td>insert before or after the $k$th element</td>
<td>$O(1)$</td>
<td>$O(n)$</td>
<td>$O(1)^1$</td>
</tr>
<tr>
<td>delete the $k$th element</td>
<td>$O(1)$</td>
<td>$O(n)$</td>
<td>$O(n)$</td>
</tr>
</tbody>
</table>

## Doubly-linked list:

<table>
<thead>
<tr>
<th>Action</th>
<th>$k = 1$</th>
<th>$1 &lt; k &lt; n$</th>
<th>$k = n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>access/change the $k$th element</td>
<td>$O(1)$</td>
<td>$O(n)$</td>
<td>$O(1)$</td>
</tr>
<tr>
<td>insert before or after the $k$th element</td>
<td>$O(1)$</td>
<td>$O(1)^1$</td>
<td>$O(1)$</td>
</tr>
<tr>
<td>delete the $k$th element</td>
<td>$O(1)$</td>
<td>$O(1)^1$</td>
<td>$O(1)$</td>
</tr>
</tbody>
</table>

$^1$ under the assumption we have a pointer to the $k$th node, $O(n)$ otherwise
• **Array indexed from 0 to \(n - 1\):**

<table>
<thead>
<tr>
<th>Operation</th>
<th>(k = 1)</th>
<th>(1 &lt; k &lt; n)</th>
<th>(k = n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access/Change the (k)th element</td>
<td>(O(1))</td>
<td>(O(n))</td>
<td>(O(1))</td>
</tr>
<tr>
<td>Insert before or after the (k)th element</td>
<td>(O(n))</td>
<td>(O(n))</td>
<td>(O(1))</td>
</tr>
<tr>
<td>Delete the (k)th element</td>
<td>(O(n))</td>
<td>(O(n))</td>
<td>(O(n))</td>
</tr>
</tbody>
</table>

• **Singly-linked list with head and tail pointers**

<table>
<thead>
<tr>
<th>Operation</th>
<th>(k = 1)</th>
<th>(1 &lt; k &lt; n)</th>
<th>(k = n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access/Change the (k)th element</td>
<td>(O(1))</td>
<td>(O(n))</td>
<td>(O(1))</td>
</tr>
<tr>
<td>Insert before or after the (k)th element</td>
<td>(O(1))</td>
<td>(O(n))</td>
<td>(O(1))</td>
</tr>
<tr>
<td>Delete the (k)th element</td>
<td>(O(1))</td>
<td>(O(n))</td>
<td>(O(n))</td>
</tr>
</tbody>
</table>

• **Doubly linked list**

<table>
<thead>
<tr>
<th>Operation</th>
<th>(k = 1)</th>
<th>(1 &lt; k &lt; n)</th>
<th>(k = n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access/Change the (k)th element</td>
<td>(O(1))</td>
<td>(O(n))</td>
<td>(O(1))</td>
</tr>
<tr>
<td>Insert before or after the (k)th element</td>
<td>(O(1))</td>
<td>(O(1))</td>
<td>(O(1))</td>
</tr>
<tr>
<td>Delete the (k)th element</td>
<td>(O(1))</td>
<td>(O(1))</td>
<td>(O(1))</td>
</tr>
</tbody>
</table>
Introduction to linked lists

• Consider the following struct definition

```c
struct node {
    string word;
    int num;
    node *next;  //pointer for the next node
};
node *p = new node;
```
Introduction to linked lists: inserting a node

- node *p;
- p = new node;
- p->num = 5;
- p->word = "Ali";
- p->next = NULL
Introduction to linked lists: adding a new node

- How can you add another node that is pointed by `p->link`?

```c
node *p;
p = new node;
p->num = 5;
p->word = "Ali";
p->next = NULL;
node *q;
```

```
p  ----> 5  Ali  
     |     |
     |     num  word  link
     v
     q
```

?
Introduction to linked lists

node *p;
p = new node;
p->num = 5;
p->word = "Ali";
p->next = NULL;

node *q;
q = new node;
Introduction to linked lists

```c
node *p, *q;
p = new node;
p->num = 5;
p->word = "Ali";
p->next = NULL;

q = new node;
q->num = 8;
q->word = "Veli";
```
Introduction to linked lists

node *p, *q;
p = new node;
p->num = 5;
p->word = "Ali";
p->next = NULL;

q = new node;
q->num=8;
q->word = "Veli";
p->next = q;
q->next = NULL;
Pointers in C/C++

```c
p = new node; delete p;
p = new node[20];

p = malloc( sizeof( node ) ); free p;

p = malloc( sizeof( node ) * 20 );
(p+10)->next = NULL; /* 11th elements */
```
Book-keeping

• malloc, new – “remember” what has been created free(p), delete (C/C++)
• When you need many small areas to be allocated, reserve a big chunk (array) and maintain your own set of free objects
• Elements of array of objects can be pointed by the pointer to an object.
Object

• Object = new object_type ;

• Equals to creating a new object with necessary size of allocated memory (delete can free it)
Some links


• **Pointer basics:**

• **C++ Memory Management : What is the difference between malloc/free and new/delete?**
Alternative: **arrays and integers**

- If you want to test pointers and linked list etc. data structures, but do not have pointers familiar (yet)

- Use arrays and indexes to array elements instead...
Replacing pointers with array index

head=3

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>next</strong></td>
<td>/</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>key</strong></td>
<td>7</td>
<td>8</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>prev</strong></td>
<td>5</td>
<td>/</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

head
Maintaining list of free objects

head=3

next | 1 | 2 | 3 | 4 | 5 | 6 | 7
--- | --- | --- | --- | --- | --- | --- | ---
/ | 5 | 1 | / | / | / | / | / 

key | 7 | 8 | 4 | / | / | / | / 
prev | 5 | / | 3 | / | / | / | / 

head=3 free=6

allocate object:
new = free;
free = next[free] ;
free object x
next[x] = free
free = x

free = -1 => array is full
Multiple lists, single free list

head1=3 => 8, 4, 7
head2=6 => 3, 9
free =2 (2)
Hack: allocate more arrays ...

use integer division and mod

\[
\text{AA}\left[\frac{i-1}{7}\right] \rightarrow \left[ (i - 1) \mod 7 \right]
\]

\[
\text{LIST}(10) = \text{AA}[1][2]
\]

\[
\text{LIST}(19) = \text{AA}[2][5]
\]
**XOR linked lists** are a data structure used in computer programming. They take advantage of the bitwise **exclusive disjunction (XOR)** operation, here denoted by $\oplus$, to decrease storage requirements for **doubly-linked lists**. An ordinary doubly-linked list stores addresses of the previous and next list items in each list node, requiring two address fields:

```
...   A   B   C   D   E   ...
    -> next  -> next  -> next  ->
    <- prev  <- prev  <- prev  <-
```

An XOR linked list compresses the same information into one address field by storing the bitwise XOR of the address for **previous** and the address for **next** in one field:

```
...   A   B   C   D   E   ...
    <- A\oplus C  <- B\oplus D  <- C\oplus E  <-
```

When you traverse the list from left to right: supposing you are at C, you can take the address of the previous item, B, and XOR it with the value in the link field (B\oplus D). You will then have the address for D and you can continue traversing the list. The same pattern applies in the other direction.
Queue (FIFO)
Queue
(basic idea, does not contain all controls!)

First = List[F]     Pop_first : \{ return List[F++] \}
Last = List[L-1]    Pop_last : \{ return List[--L] \}

Full: return ( L == MAX_SIZE )
Empty: F < 0 or F >= L
A circular buffer or ring buffer is a data structure that uses a single, fixed-size buffer as if it were connected end-to-end. This structure lends itself easily to buffering data streams.
Circular Queue

First = List[F]

**Add_to_end** (x) : { List[L] = x ; L = (L+1) % MAX_SIZE }  // % = modulo

Last = List[ (L-1+MAX_SIZE) % MAX_SIZE ]

Full: return ( (L+1)%MAX_SIZE == F )

Empty: F==L
Queue

- enqueue(x) - add to end
- dequeue() - fetch from beginning

- FIFO – First In First Out

- $O(1)$ in all reasonable cases 😊
Stack

• push(x) -- add to end (add to top)
• pop() -- fetch from end (top)

• O(1) in all reasonable cases 😊

• LIFO – Last In, First Out
Stack based languages

• Implement a postfix calculator
  – Reverse Polish notation

• \(5 4 3 \times 2 - +\) => \(5+((4\times3)-2)\)

• Very simple to \textit{parse} and interpret

• \textit{FORTH, Postscript} are stack-based languages
Array based stack

• How to know how big a stack shall be?

• When full, allocate bigger table dynamically, and copy all previous values there

• $O(n)$ ?
• When full, create 2x bigger table, copy previous n elements:

• After every $2^k$ insertions, perform $O(n)$ copy

• $O(n)$ individual insertions +
• $n/2 + n/4 + n/8$ ... copy-ing
• Total: $O(n)$ effort!
• when n=32 -> 33  (copy 32, insert 1)
• delete: 33->32
  – should you delete immediately?
  – Delete only when becomes less than 1/4th full
  – Have to delete at least n/2 to decrease
  – Have to add at least n to increase size
  – Most operations, $O(1)$ effort
  – But few operations take $O(n)$ to copy
  – For any $m$ operations, $O(m)$ time
Amortized analysis

• Analyze the time complexity over the entire “lifespan” of the algorithm

• Some operations that cost more will be “covered” by many other operations taking less
Lists and **dictionary ADT**...

- How to maintain a dictionary using (linked) lists?
- **Is k in D?**
  - go through all elements d of D, test if d==k  \(O(n)\)
  - If sorted: d= first(D); while( d<=k ) d=next(D);
  - on average  \(n/2\) tests ...
- **Add(k,D) => insert(k,D) = O(1)  or O(n) – test for uniqueness**
Array based sorted list

• is \( d \) in \( D \)?
• Binary search in \( D \)

![Diagram of binary search in a sorted list with low, mid, and high markers.]
Binary search / recursive

BinarySearch(A[0..N-1], value, low, high)
{
    if (high < low)
        return -1 // not found
    mid = low + ((high - low) / 2)  // Note: not (low + high) / 2 !!
    if (A[mid] > value)
        return BinarySearch(A, value, low, mid-1)
    else if (A[mid] < value)
        return BinarySearch(A, value, mid+1, high)
    else
        return mid // found
}
Binary search – Iterative

**BinarySearch**(A[0..N-1], value)
{
    low = 0 ; high = N - 1 ;
    while (low <= high) {
        mid = low + ((high - low) / 2) // Note: not (low + high) / 2 !!
        if (A[mid] > value)
            high = mid - 1
        else if (A[mid] < value)
            low = mid + 1
        else
            return mid // found
    }
    return -1 // not found
}
Work performed

- \( x \equiv A[18] \ ? \ < \)
- \( x \equiv A[9] \ ? \ > \)
- \( x \equiv A[13] \ ? \ == \)
- \( O(\lg n) \)
Sorting

• given a list, arrange values so that
  \[ L[1] \leq L[2] \leq ... \leq L[n] \]
• \( n \) elements => \( n! \) possible orderings
• One test \( L[i] \leq L[j] \) can divide \( n! \) to 2
  – Make a binary tree and calculate the depth
• \( \log(n!) = \Omega(n \log n) \)
• Hence, lower bound for sorting is \( \Omega(n \log n) \)
  – using comparisons...
Proof: $\log(n!) = \Omega \left( n \log n \right)$

- $\log(n!) = \log n + \log(n-1) + \log(n-2) + \ldots + \log(1)$
  
  $\geq \frac{n}{2} \log \left( \frac{n}{2} \right)$

  $= \Omega \left( n \log n \right)$
Decision-tree example

Sort $\langle a_1, a_2, a_3 \rangle$

$\langle 9, 4, 6 \rangle$:

```
   1:2
  /   \
2:3   1:3
 /     / \
123  1:3  213
   /       /\
  132  312  231
   /             \\
  321
```

$4 \leq 6 \leq 9$

Each leaf contains a permutation $\langle \pi(1), \pi(2), \ldots, \pi(n) \rangle$ to indicate that the ordering $a_{\pi(1)} \leq a_{\pi(2)} \leq \cdots \leq a_{\pi(n)}$ has been established.
Decision tree model

- $n!$ orderings (leaves)
- Height of such tree?
\[ [\log_2 n!] \geq \log_2 n! \geq \sum_{i=1}^{n} \log_2 i \geq \sum_{i=1}^{n/2} \log_2 \frac{n}{2} \geq \frac{n}{2} \log_2 \frac{n}{2} = \Omega(n \log n). \]
• \( \log(n!) = \log(n) + \log(n-1) + \ldots + \log(1) \)

a) \( \leq n \log(n) \)

b) \( \geq n/2 \cdot \log(n/2) = n/2 \log n - n/2 \)
Lower bound for decision-tree sorting

**Theorem.** Any decision tree that can sort $n$ elements must have height $\Omega(n \lg n)$.

**Proof.** The tree must contain $\geq n!$ leaves, since there are $n!$ possible permutations. A height-$h$ binary tree has $\leq 2^h$ leaves. Thus, $n! \leq 2^h$.

\[ h \geq \lg(n!) \geq \lg \left( \left(\frac{n}{e}\right)^n \right) = n \lg n - n \lg e = \Omega(n \lg n). \]
The divide-and-conquer design paradigm

1. *Divide* the problem (instance) into subproblems.

2. *Conquer* the subproblems by solving them recursively.

3. *Combine* subproblem solutions.
Merge sort

\textbf{Merge-Sort}(A, p, r)

if \( p < r \)

then \( q = \lfloor \frac{p+r}{2} \rfloor \)

\textbf{Merge-Sort}( A, p, q )

\textbf{Merge-Sort}( A, q+1, r )

\textbf{Merge}( A, p, q, r )

It was invented by \textit{John von Neumann} in 1945.
Example

• Applying the merge sort algorithm:

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```

```
13 77 49 35 61 48 3 23 95 73 89 37 57 99 94 28 15 55 7 51 88 97 62
```
Merge of two lists: \( \Theta(n) \)

A, B – lists to be merged

L = new list; // empty

while( A not empty and B not empty )
    if A.first() \leq B.first()
        then append( L, A.first() ); A = rest(A);
    else append( L, B.first() ); B = rest(B);

append( L, A); // all remaining elements of A
append( L, B ); // all remaining elements of B

return L
Wikipedia / viz.

Value

Pos in array
Run-time Analysis of Merge Sort

• Thus, the time required to sort an array of size $n > 1$ is:
  – the time required to sort the first half,
  – the time required to sort the second half, and
  – the time required to merge the two lists

• That is:

$$T(n) = \begin{cases} 
\Theta(1) & n = 1 \\
2T\left(\frac{n}{2}\right) + \Theta(n) & n > 1 
\end{cases}$$
Recursion tree

Solve $T(n) = 2T(n/2) + cn$, where $c > 0$ is constant.

$h = \log n$

$\Theta(1)$

$\#leaves = n$

Total = $\Theta(n \log n)$
Merge sort

• Worst case, average case, best case ...

\[ \Theta( n \log n ) \]

• Common wisdom:
  – Requires additional space for merging (in case of arrays)

• Homework*: develop in-place merge of two lists implemented in arrays /compare speed/

\[ a \]
\[ L[a] \leq L[b] \]
\[ L[a] > L[b] \]
Quicksort

- Divide-and-conquer algorithm.
- Sorts “in place” (like insertion sort, but not like merge sort).
- Very practical (with tuning).
Divide and conquer

Quicksort an $n$-element array:

1. **Divide**: Partition the array into two subarrays around a **pivot** $x$ such that elements in lower subarray $\leq x \leq$ elements in upper subarray.

   \[ \leq x \quad x \quad \geq x \]

2. **Conquer**: Recursively sort the two subarrays.

3. **Combine**: Trivial.

   **Key**: Linear-time partitioning subroutine.
Pseudocode for quicksort

\[
\text{QUICKSORT}(A, p, r) \\
\text{if } p < r \\
\quad \text{then } q \leftarrow \text{PARTITION}(A, p, r) \\
\quad \text{QUICKSORT}(A, p, q-1) \\
\quad \text{QUICKSORT}(A, q+1, r)
\]

Initial call: \text{QUICKSORT}(A, 1, n)
Partitioning subroutine

\textsc{Partition}(A, p, q) \triangleright A[p \ldots q]

\begin{align*}
x & \leftarrow A[p] \quad \triangleright pivot = A[p] \\
i & \leftarrow p \\
\text{for } j & \leftarrow p + 1 \text{ to } q \\
\text{do if } A[j] & \leq x \\
\text{then } & i \leftarrow i + 1 \\
\text{exchange } & A[i] \leftrightarrow A[j] \\
\text{exchange } & A[p] \leftrightarrow A[i] \\
\text{return } & i
\end{align*}

\textbf{Invariant:} \quad \begin{bmatrix} x & \leq x & \geq x & ? \\ p & i & j & q \end{bmatrix}

\text{Running time } = O(n) \text{ for } n \text{ elements.}
pivot = A[R];  //
i=L; j=R-1;
while( i<=j )
    while ( A[i] < pivot ) i++ ;  // will stop at pivot latest
    while ( i<=j and A[j] >=  pivot ) j-- ;
A[R]=A[i];
A[i]=pivot;
return i;
Wikipedia / “video”
Worst-case of quicksort

- Input sorted or reverse sorted.
- Partition around min or max element.
- One side of partition always has no elements.

\[ T(n) = T(0) + T(n-1) + \Theta(n) \]
\[ = \Theta(1) + T(n-1) + \Theta(n) \]
\[ = T(n-1) + \Theta(n) \]
\[ = \Theta(n^2) \quad \text{(arithmetic series)} \]
Best-case analysis
(For intuition only!)

If we’re lucky, \textsc{Partition} splits the array evenly:

\[
T(n) = 2T(n/2) + \Theta(n) = \Theta(n \log n) \quad \text{(same as merge sort)}
\]

What if the split is always $\frac{1}{10} : \frac{9}{10}$?

\[
T(n) = T\left(\frac{1}{10} n\right) + T\left(\frac{9}{10} n\right) + \Theta(n)
\]

What is the solution to this recurrence?
Analysis of “almost-best” case

\[
T\left(\frac{1}{10} n\right) \quad c n \quad T\left(\frac{9}{10} n\right)
\]
Analysis of “almost-best” case

\[ \log_{10} n \]

\[ \frac{1}{10} cn \]

\[ \frac{9}{100} cn \]

\[ \frac{9}{100} cn \]

\[ \frac{81}{100} cn \]

\[ O(n) \text{ leaves} \]

\[ \Theta(1) \]

\[ \Theta(n \log n) \]

Lucky!

\[ cn \log_{10} n \leq T(n) \leq cn \log_{10/9} n + O(n) \]
More intuition

Suppose we alternate lucky, unlucky, lucky, unlucky, lucky, ....

\[ L(n) = 2U(n/2) + \Theta(n) \quad \text{lucky} \]
\[ U(n) = L(n - 1) + \Theta(n) \quad \text{unlucky} \]

Solving:

\[ L(n) = 2(L(n/2 - 1) + \Theta(n/2)) + \Theta(n) \]
\[ = 2L(n/2 - 1) + \Theta(n) \]
\[ = \Theta(n \lg n) \quad \text{Lucky!} \]

How can we make sure we are usually lucky?
Choice of pivot

- Select median of three ...

- Select random – opponent can not choose the winning strategy against you!
Randomized quicksort

**IDEA**: Partition around a *random* element.

- Running time is independent of the input order.
- No assumptions need to be made about the input distribution.
- No specific input elicits the worst-case behavior.
- The worst case is determined only by the output of a random-number generator.
Random pivot

Select pivot randomly from the region (blue) and swap with last position
Select pivot as a median of 3 [or more] random values from region

Apply non-recursive sort for array less than 10-20
Randomized quicksort analysis

Let $T(n)$ = the random variable for the running time of randomized quicksort on an input of size $n$, assuming random numbers are independent.

For $k = 0, 1, \ldots, n-1$, define the indicator random variable

$$X_k = \begin{cases} 
1 & \text{if PARTITION generates a } k : n-k-1 \text{ split,} \\
0 & \text{otherwise.}
\end{cases}$$

$$E[X_k] = \Pr\{X_k = 1\} = 1/n,$$ since all splits are equally likely, assuming elements are distinct.
Analysis (continued)

\[ T(n) = \begin{cases} 
T(0) + T(n-1) + \Theta(n) & \text{if } 0:n-1 \text{ split,} \\
T(1) + T(n-2) + \Theta(n) & \text{if } 1:n-2 \text{ split,} \\
\vdots \\
T(n-1) + T(0) + \Theta(n) & \text{if } n-1:0 \text{ split,} 
\end{cases} \]

\[ = \sum_{k=0}^{n-1} X_k \left( T(k) + T(n-k-1) + \Theta(n) \right) \]
Calculating expectation

\[ E[T(n)] = E \left[ \sum_{k=0}^{n-1} X_k (T(k) + T(n-k-1) + \Theta(n)) \right] \]

Take expectations of both sides.
Calculating expectation

\[
E[T(n)] = E\left[ \sum_{k=0}^{n-1} X_k (T(k) + T(n - k - 1) + \Theta(n)) \right]
\]

\[
= \sum_{k=0}^{n-1} E[X_k (T(k) + T(n - k - 1) + \Theta(n))]
\]

Linearity of expectation.
Calculating expectation

\[ E[T(n)] = E \left[ \sum_{k=0}^{n-1} X_k (T(k) + T(n-k-1) + \Theta(n)) \right] \]

\[ = \sum_{k=0}^{n-1} E[X_k] \cdot E[T(k) + T(n-k-1) + \Theta(n)] \]

Independence of \( X_k \) from other random choices.
Calculating expectation

\[ E[T(n)] = E \left[ \sum_{k=0}^{n-1} X_k (T(k) + T(n - k - 1) + \Theta(n)) \right] \]

\[ = \sum_{k=0}^{n-1} E[X_k (T(k) + T(n - k - 1) + \Theta(n))] \]

\[ = \sum_{k=0}^{n-1} E[X_k] \cdot E[T(k) + T(n - k - 1) + \Theta(n)] \]

\[ = \frac{1}{n} \sum_{k=0}^{n-1} E[T(k)] + \frac{1}{n} \sum_{k=0}^{n-1} E[T(n - k - 1)] + \frac{1}{n} \sum_{k=0}^{n-1} \Theta(n) \]

Linearity of expectation; \( E[X_k] = 1/n \).
Calculating expectation

\[ E[T(n)] = E \left[ \sum_{k=0}^{n-1} X_k (T(k) + T(n - k - 1) + \Theta(n)) \right] \]

\[ = \sum_{k=0}^{n-1} E[X_k] \cdot E[T(k) + T(n - k - 1) + \Theta(n)] \]

\[ = \frac{1}{n} \sum_{k=0}^{n-1} E[T(k)] + \frac{1}{n} \sum_{k=0}^{n-1} E[T(n - k - 1)] + \frac{1}{n} \sum_{k=0}^{n-1} \Theta(n) \]

\[ = \frac{2}{n} \sum_{k=1}^{n-1} E[T(k)] + \Theta(n) \]

Summations have identical terms.
Hairy recurrence

\[ E[T(n)] = 2 \sum_{k=2}^{n-1} E[T(k)] + \Theta(n) \]

(The \( k = 0, 1 \) terms can be absorbed in the \( \Theta(n) \).

Prove: \( E[T(n)] \leq an \log n \) for constant \( a > 0 \).

- Choose \( a \) large enough so that \( an \log n \) dominates \( E[T(n)] \) for sufficiently small \( n \geq 2 \).

Use fact: \( \sum_{k=2}^{n-1} k \log k \leq \frac{1}{2} n^2 \log n - \frac{1}{8} n^2 \) (exercise).
Substitution method

\[ E[T(n)] \leq \frac{2}{n} \sum_{k=2}^{n-1} ak \log k + \Theta(n) \]

\[ = \frac{2a}{n} \left( \frac{1}{2} n^2 \log n - \frac{1}{8} n^2 \right) + \Theta(n) \]

\[ = an \log n - \left( \frac{an}{4} - \Theta(n) \right) \]

\[ \leq an \log n , \]

if \( a \) is chosen large enough so that \( an/4 \) dominates the \( \Theta(n) \).
Quicksort in practice

- Quicksort is a great general-purpose sorting algorithm.
- Quicksort is typically over twice as fast as merge sort.
- Quicksort can benefit substantially from *code tuning*.
- Quicksort behaves well even with caching and virtual memory.
We can sort in $O(n \log n)$

• Is that the best we can do?

• Remember: using comparisons $<$, $>$, $\leq$, $\geq$ we cannot do better than $O(n \log n)$
How fast can we sort \( n \) integers?

- E.g. sort people by year of birth?
- Sort people by sex?
Sorting in linear time

Counting sort: No comparisons between elements.

- **Input**: $A[1 \ldots n]$, where $A[j] \in \{1, 2, \ldots, k\}$.
- **Output**: $B[1 \ldots n]$, sorted.
- **Auxiliary storage**: $C[1 \ldots k]$. 

Counting sort

for $i \leftarrow 1$ to $k$
    do $C[i] \leftarrow 0$

for $j \leftarrow 1$ to $n$
    do $C[A[j]] \leftarrow C[A[j]] + 1$ \>
    \> $C[i] = |\{\text{key} = i\}|$

for $i \leftarrow 2$ to $k$
    do $C[i] \leftarrow C[i] + C[i-1]$ \>
    \> $C[i] = |\{\text{key} \leq i\}|$

for $j \leftarrow n$ downto $1$
    do $B[C[A[j]]] \leftarrow A[j]$
    \> $C[A[j]] \leftarrow C[A[j]] - 1$
Loop 1

\[ A: \quad 4 \quad 1 \quad 3 \quad 4 \quad 3 \]

\[ B: \quad \]

\[ C: \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \]

\[ \text{for } i \leftarrow 1 \text{ to } k \]
\[ \text{do } C[i] \leftarrow 0 \]

September 26, 2005
Copyright © 2001-5 Erik D. Demaine and Charles E. Leiserson
Loop 2

\[ A: \begin{array}{ccccc}
4 & 1 & 3 & 4 & 3 \\
\end{array} \]

\[ C: \begin{array}{ccccc}
1 & 2 & 3 & 4 & 2 \\
\end{array} \]

\[ B: \begin{array}{cccc}
& & & \\
\end{array} \]

\[ \text{for } j \leftarrow 1 \text{ to } n \]
\[ \text{do } C[A[j]] \leftarrow C[A[j]] + 1 \quad \triangleright \quad C[i] = |\{\text{key} = i\}| \]
Loop 3

\[ A: \begin{array}{cccccc}
4 & 1 & 3 & 4 & 3 \\
\end{array} \]

\[ C: \begin{array}{cccc}
1 & 0 & 2 & 2 \\
\end{array} \]

\[ B: \begin{array}{cccccc}
\phantom{1} & \phantom{1} & \phantom{1} & \phantom{1} & \phantom{1} & \phantom{1} \\
\end{array} \]

\[ C': \begin{array}{cccc}
1 & 1 & 3 & 5 \\
\end{array} \]

\[ \text{for } i \leftarrow 2 \text{ to } k \]
\[ \text{do } C[i] \leftarrow C[i] + C[i-1] \quad \triangledown C[i] = \left| \{\text{key } \leq i\} \right| \]
Loop 4

A: 4 1 3 4 3
B: 3 4

C: 1 1 2 5
C': 1 1 2 4

for j ← n downto 1
    do B[C[A[j]]] ← A[j]
       C[A[j]] ← C[A[j]] - 1
Analysis

$$\Theta(k)$$

\[
\begin{align*}
\text{for } i & \leftarrow 1 \text{ to } k \\
\text{do } C[i] & \leftarrow 0
\end{align*}
\]

$$\Theta(n)$$

\[
\begin{align*}
\text{for } j & \leftarrow 1 \text{ to } n \\
\text{do } C[A[j]] & \leftarrow C[A[j]] + 1
\end{align*}
\]

$$\Theta(k)$$

\[
\begin{align*}
\text{for } i & \leftarrow 2 \text{ to } k \\
\text{do } C[i] & \leftarrow C[i] + C[i-1]
\end{align*}
\]

$$\Theta(n)$$

\[
\begin{align*}
\text{for } j & \leftarrow n \text{ downto } 1 \\
\text{do } B[C[A[j]]] & \leftarrow A[j] \\
C[A[j]] & \leftarrow C[A[j]] - 1
\end{align*}
\]

$$\Theta(n + k)$$
Running time

If $k = O(n)$, then counting sort takes $\Theta(n)$ time.

- But, sorting takes $\Omega(n \log n)$ time!
- Where’s the fallacy?

Answer:

- *Comparison sorting* takes $\Omega(n \log n)$ time.
- Counting sort is not a *comparison sort*.
- In fact, not a single comparison between elements occurs!
Stable sorting

Counting sort is a *stable* sort: it preserves the input order among equal elements.

Exercise: What other sorts have this property?
Radix sort

- **Origin**: Herman Hollerith’s card-sorting machine for the 1890 U.S. Census. (See Appendix 1.)
- Digit-by-digit sort.
- Hollerith’s original (bad) idea: sort on most-significant digit first.
- Good idea: Sort on least-significant digit first with auxiliary stable sort.
Radix sort

Radix-Sort(A,d)

1. for i = 1 to d /* least significant to most significant */
2. use a stable sort to sort A on digit i
Operation of radix sort

```
329  720  720  329
457  355  329  355
657  436  436  436
839  457  839  457
436  657  355  657
720  329  457  720
355  839  657  839
```
Correctness of radix sort

**Induction on digit position**

- Assume that the numbers are sorted by their low-order $t - 1$ digits.

- Sort on digit $t$
  - Two numbers that differ in digit $t$ are correctly sorted.
Correctness of radix sort

*Induction on digit position*

- Assume that the numbers are sorted by their low-order \( t - 1 \) digits.

- Sort on digit \( t \)
  - Two numbers that differ in digit \( t \) are correctly sorted.
  - Two numbers equal in digit \( t \) are put in the same order as the input \( \Rightarrow \) correct order.
Analysis of radix sort

- Assume counting sort is the auxiliary stable sort.
- Sort $n$ computer words of $b$ bits each.
- Each word can be viewed as having $b/r$ base-$2^r$ digits.

**Example:** 32-bit word

$r = 8 \Rightarrow b/r = 4$ passes of counting sort on base-$2^8$ digits; or $r = 16 \Rightarrow b/r = 2$ passes of counting sort on base-$2^{16}$ digits.

*How many passes should we make?*
Analysis (continued)

**Recall:** Counting sort takes $\Theta(n + k)$ time to sort $n$ numbers in the range from 0 to $k - 1$.

If each $b$-bit word is broken into $r$-bit pieces, each pass of counting sort takes $\Theta(n + 2^r)$ time. Since there are $b/r$ passes, we have

$$T(n, b) = \Theta\left(\frac{b}{r} \left(n + 2^r\right)\right).$$

Choose $r$ to minimize $T(n, b)$:

- Increasing $r$ means fewer passes, but as $r \gg \log n$, the time grows exponentially.
Choosing $r$

$$T(n, b) = \Theta\left(\frac{b}{r} (n + 2^r)\right)$$

Minimize $T(n, b)$ by differentiating and setting to 0.

Or, just observe that we don’t want $2^r \gg n$, and there’s no harm asymptotically in choosing $r$ as large as possible subject to this constraint.

Choosing $r = \lg n$ implies $T(n, b) = \Theta(bn/\lg n)$.

- For numbers in the range from 0 to $n^d - 1$, we have $b = d \lg n \Rightarrow$ radix sort runs in $\Theta(d n)$ time.
Conclusions

In practice, radix sort is fast for large inputs, as well as simple to code and maintain.

**Example** (32-bit numbers):
- At most 3 passes when sorting $\geq 2000$ numbers.
- Merge sort and quicksort do at least $\lceil \lg 2000 \rceil = 11$ passes.

**Downside:** Unlike quicksort, radix sort displays little locality of reference, and thus a well-tuned quicksort fares better on modern processors, which feature steep memory hierarchies.
Radix sort using lists (stable)
Radix sort using lists (stable)
Radix sort using lists (stable)

1. a → bba → aba → cca
   b → bbb → adb → ccb
   c → aac → ccc
   d → aad

2. a → aac → aad
   b → bba → aba → bbb
   c → cca → ccb → ccc
   d → adb

3. a → aac → aad → aba → adb
   b → bba → bbb
   c → cca → ccb → ccc
Why not from left to right?

0101100  0101100  0101100  0101100
1001010  0010010  0010010  0010010
1111000  1111000  0101000  0101000
1001001  1001001  1001001  0010000
0010010  1001010  1001010  1001010
1001001  1001001  1001001  1001001
0101000  0101000  1111000  1111000
0010000  0010000  0010000  1001001

- Swap ‘0’ with first ‘1’
- Idea 1: recursively sort first and second half
  - Exercise?
Bitwise sort left to right

• Idea2:
  – swap elements only if the prefixes match...

  – For all bits from most significant
    • advance when 0
    • when 1 -> look for next 0
      – if prefix matches, swap
      – otherwise keep advancing on 0’s and look for next 1
Bitwise left to right sort

;/* Historical sorting – was used in Univ. of Tartu using assembler…. */
/* C implementation – Jaak Vilo, 1989 */

void bitwisesort( SORTTYPE *ARRAY , int size )
{
  int i, j, tmp, nrbits ;

  register SORTTYPE mask , curbit , group ;

  nrbits = sizeof( SORTTYPE ) * 8 ;

  curbit = 1 << (nrbits-1) ;  /* set most significant bit 1 */
  mask   = 0;               /* mask of the already sorted area */

  Jaak Vilo, Univ. of Tartu
do {  /* For each bit */
    i=0;

new_mask:  
    for( ; ( i < size ) & ( ! (ARRAY[i] & curbit) ) ; i++ )  ; /* Advance while bit == 0 */
    if( i >= size ) goto array_end ;
    group = ARRAY[i] & mask ;  /* Save current prefix snapshot */

    j=i;  /* memorize location of 1 */
    for( ; ; ) {
      if ( ++i >= size ) goto array_end ;  /* reached end of array */
      if ( (ARRAY[i] & mask) != group ) goto new_mask ;  /* new prefix */

      if ( ! (ARRAY[i] & curbit) ) { /* bit is 0 – need to swap with previous location of 1, A[i] ↔ A[j] */
        tmp = ARRAY[i];  ARRAY[i] = ARRAY[j];  ARRAY[j] = tmp ;  j += 1 ; /* swap and increase j to the next possible 1 */
      }
    }
}
array_end:
    mask = mask | curbit ;  /* area under mask is now sorted */
    curbit >>= 1 ;  /* next bit */
} while( curbit );  /* until all bits have been sorted... */

Jaak Vilo, Univ. of Tartu
Bitwise from left to right

0010000
0010010
0101000
0101100
1001010
1001001
1001001
1111000

• Swap ‘0’ with first ‘1’
Bucket sort

• Assume uniform distribution

• Allocate $O(n)$ buckets

• Assign each value to pre-assigned bucket
Sort small buckets with insertion sort
http://sortbenchmark.org/

- Minutesort – max amount sorted in 1 minute
  - 116GB in 58.7 sec (Jim Wyllie, IBM Research)
  - 40-node 80-Itanium cluster, SAN array of 2,520 disks
- 2009, **500 GB Hadoop** 1406 nodes x (2 Quadcore Xeons, 8 GB memory, 4 SATA) Owen O'Malley and Arun Murthy Yahoo Inc.
- Performance / Price Sort and PennySort
# Year 2008 Results

**commentary by Mehul Shah on 2007 winners**

<table>
<thead>
<tr>
<th>Penny</th>
<th>Daytona</th>
<th>Indy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>(new 2008)</em> 1,812 M records (181 GB) in 2,408 seconds</td>
<td><em>(new 2008)</em> 1,896 M records (190 GB) in 2,408 seconds</td>
</tr>
<tr>
<td></td>
<td><code>psort</code></td>
<td><code>psort</code></td>
</tr>
<tr>
<td></td>
<td>2.4 Ghz AMD Athlon 64, 2 GB RAM, 4x160GB SATA disks, Linux</td>
<td>2.4 Ghz AMD Athlon 64, 2 GB RAM, 4x160GB SATA disks, Linux</td>
</tr>
<tr>
<td></td>
<td>Paolo Bertasi, Marco Bressan and Enoch Peserico</td>
<td>Paolo Bertasi, Marco Bressan and Enoch Peserico</td>
</tr>
<tr>
<td></td>
<td>Univ. Padova, Italy</td>
<td>Univ. Padova, Italy</td>
</tr>
</tbody>
</table>

| Minute         | 214 GB (2140 million records) **TokuSampleSort** tx2500 disk cluster, 400 nodes x (2 processors, 6-disk RAID, 8 GB memory) **Bradley C. Kuszmaul, MIT** | 264 GB (2640 M records) **TokuSampleSort** tx2500 disk cluster, 400 nodes x (2 processors, 6-disk RAID, 8 GB memory) **Bradley C. Kuszmaul, MIT** |

| TeraByte       | *(new 2008)* 209 seconds (3.48 minutes) **Hadoop** 910 nodes x (4 dual-core processors, 4 disks, 8 GB memory) Owen O’Malley, Yahoo | 197 seconds (3.28 minutes) **TokuSampleSort** tx2500 disk cluster, 400 nodes x (2 processors, 6-disk RAID, 8 GB memory) **Bradley C. Kuszmaul, MIT** |

| Joule          | *(2007)* 10 GB sorted using 8.6 kJoules 11,600 records sorted / joule **CoolSort** Mobile Core 2 Duo, 13 SATA laptop disks, **Nsort Suzanne Rivoire** (Stanford), **Mehul A. Shah** (HP Labs), **Partha Ranganathan** (HP Labs), **Christos Kozyrakis** (Stanford) | |

| Joule          | *(2007)* 100 GB sorted using 88 kJoules | |
## Year 2009 Results

<table>
<thead>
<tr>
<th>Gray</th>
<th>Dayton 2009, 0.578 TB/min</th>
<th>Indy 2009, 0.564 TB/min</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hadoop</td>
<td>DEMSort</td>
</tr>
<tr>
<td></td>
<td>100 TB in 173 minutes</td>
<td>1,000,029,388,800 records in 10,628 seconds</td>
</tr>
<tr>
<td></td>
<td>3452 nodes x (2 Quadcore Xeons, 8 GB memory, 4 SATA)</td>
<td>195 nodes x (2 Quadcore processors, 16 GB memory, 4x250GB disks)</td>
</tr>
<tr>
<td></td>
<td>Owen O’Malley and Arun Murthy</td>
<td>Mirko Rahn, Peter Sanders, Johannes Singler and Tim Kieritz</td>
</tr>
<tr>
<td></td>
<td>Yahoo Inc.</td>
<td>Karlsruhe Institute of Technology, Germany</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Penny</th>
<th>2009, 223 GB</th>
<th>2-way tie: 2009, 246 GB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>psort</td>
<td>OzSort</td>
</tr>
<tr>
<td></td>
<td>2.6 Ghz AMD Athlon LE 1640, 4 GB RAM, 5x160 GB 7200 RPM SATA, Linux</td>
<td>2.7 Ghz AMD Kuna X2 7750+, 4 GB RAM, 5x160 GB 7200 RPM SATA, Linux</td>
</tr>
<tr>
<td></td>
<td>Paolo Bertasi, Marco Bressan and Enoch Peserico</td>
<td>Nikolas Asktis and Ranjan Sinha</td>
</tr>
<tr>
<td></td>
<td>Univ. Padova, Italy</td>
<td>Univ. Melbourne, Australia</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Minute</th>
<th>2009, 500 GB</th>
<th>2009, 955 GB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hadoop</td>
<td>DEMSort</td>
</tr>
<tr>
<td></td>
<td>1406 nodes x (2 Quadcore Xeons, 8 GB memory, 4 SATA)</td>
<td>195 nodes x (2 Quadcore processors, 16 GB memory, 4x250GB disks)</td>
</tr>
<tr>
<td></td>
<td>Owen O’Malley and Arun Murthy</td>
<td>Mirko Rahn, Peter Sanders, Johannes Singler and Tim Kieritz</td>
</tr>
<tr>
<td></td>
<td>Yahoo Inc.</td>
<td>Karlsruhe Institute of Technology, Germany</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Joule 10^6 recs</th>
<th>2007, 8.6 kJoules</th>
<th>2007, 88 kJoules</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoolSort</td>
<td>11,600 records sorted / joule</td>
<td>11,600 records sorted / joule</td>
</tr>
<tr>
<td></td>
<td>Mobile Core 2 Duo, 13 SATA laptop disks, Nsort</td>
<td>Mobile Core 2 Duo, 13 SATA laptop disks, Nsort</td>
</tr>
<tr>
<td></td>
<td>Suzanne Rivoire (Stanford), Mehul A. Shah (HP Labs), Partha Ranganathan (HP Labs), Christos Kozyrakis (Stanford)</td>
<td>Suzanne Rivoire (Stanford), Mehul A. Shah (HP Labs), Partha Ranganathan (HP Labs), Christos Kozyrakis (Stanford)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Joule 10^8 recs</th>
<th>2009, 87 kJoules</th>
<th>2009, 87 kJoules</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sort Benchmark

• [http://sortbenchmark.org/](http://sortbenchmark.org/)
• Sort Benchmark Home Page
• We have a new benchmark called new GraySort, new in memory of the father of the sort benchmarks, Jim Gray. It replaces TeraByteSort which is now retired.
• Unlike 2010, we will not be accepting early entries for the 2011 year. The deadline for submitting entries is April 1, 2011.
  – All hardware used must be off-the-shelf and unmodified.
  – For Daytona cluster sorts where input sampling is used to determine the output partition boundaries, the input sampling must be done evenly across all input partitions.

**New rules for GraySort:**
• The input file size is now minimum ~100TB or 1T records. Entries with larger input sizes also qualify.
• The winner will have the fastest SortedRecs/Min.
• We now provide a new input generator that works in parallel and generates binary data. See below.
• For the Daytona category, we have two new requirements. (1) The sort must run continuously (repeatedly) for a minimum 1 hour. (This is a minimum reliability requirement). (2) The system cannot overwrite the input file.
Order statistics

• Minimum – the smallest value
• Maximum – the largest value
• In general i’th value.
• Find the median of the values in the array
• Median in sorted array A :
  – n is odd  $A[(n+1)/2]$
  – n is even – $A\lceil (n+1)/2 \rceil$ or $A\lfloor (n+1)/2 \rfloor$
Order statistics

• Input: A set $A$ of $n$ numbers and $i$, $1 \leq i \leq n$
• Output: $x$ from $A$ that is larger than exactly $i-1$ elements of $A$
Minimum

Minimum(A)

1 \text{ min } = A[1]

2 \text{ for } i = 2 \text{ to } \text{length}(A)

3 \quad \text{if } \text{min} > A[i]

4 \quad \text{then } \text{min} = A[i]

5 \text{ return min}

n-1 comparisons.
Min and max together

• compare every two elements $A[i], A[i+1]$  
• Compare larger against current max  
• Smaller against current min

• $\lceil n / 2 \rceil$
Selection in expected $O(n)$

Randomised-select( $A$, $p$, $r$, $i$ )

if $p=r$ then return $A[p]$

$q =$ Randomised-Partition($A$, $p$, $r$)

$k = q - p + 1$ // nr of elements in subarr

if $i \leq k$

then return Randomised-Partition($A$, $p$, $q$, $i$)

else return Randomised-Partition($A$, $q+1$, $r$, $i-k$)
Conclusion

- Sorting in general $O(n \log n)$
- Quicksort is rather good

- Linear time sorting is achievable when one does not assume only direct comparisons

- Find $i$’th value – expected $O(n)$

- Find $i$’th value: worst case $O(n)$ – see CLRS
Lists: Array

- Array:
  
  Before:
  0 1
  3 6 7 5 2

  Insert 8 after L[2]:
  0 1
  3 6 7 8 5 2

  Delete last:
  0 1
  3 6 7 8 5 2
Ok...

• lists – a versatile data structure for various purposes
• Sorting – a typical algorithm (many ways)
• Which sorting methods for array/list?

• Array: most of the important (e.g. update) tasks seem to be O(n), which is bad
• **Array indexed from 0 to** \( n - 1 \):

<table>
<thead>
<tr>
<th></th>
<th>( k = 1 )</th>
<th>( 1 &lt; k &lt; n )</th>
<th>( k = n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>access/change the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(1) )</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>insert before or after the ( k )th element</td>
<td>( O(n) )</td>
<td>( O(n) )</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>delete the ( k )th element</td>
<td>( O(n) )</td>
<td>( O(n) )</td>
<td>( O(1) )</td>
</tr>
</tbody>
</table>

• **Singly-linked list with head and tail pointers**

<table>
<thead>
<tr>
<th></th>
<th>( k = 1 )</th>
<th>( 1 &lt; k &lt; n )</th>
<th>( k = n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>access/change the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(n) )</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>insert before or after the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(n) )</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>delete the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(n) )</td>
<td>( O(n) )</td>
</tr>
</tbody>
</table>

• **Doubly linked list**

<table>
<thead>
<tr>
<th></th>
<th>( k = 1 )</th>
<th>( 1 &lt; k &lt; n )</th>
<th>( k = n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>access/change the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(n) )</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>insert before or after the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(1) )</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>delete the ( k )th element</td>
<td>( O(1) )</td>
<td>( O(1) )</td>
<td>( O(1) )</td>
</tr>
</tbody>
</table>
Can we search faster in linked lists?

• Why sort linked lists if search anyway $O(n)$?

• Linked lists:
  – what is the “mid-point” of any sublist?
  – Therefore, binary search can not be used...

• Or can it?
Skip List

A skip list, introduced by Pugh [Pugh 1990], is a randomized balanced tree data structure organized as a tower of increasingly sparse linked lists. Level 0 of a skip list is a linked list of all nodes in increasing order by key. For each $i$ greater than 0, each node in level $i-1$ appears in level $i$ independently with some fixed probability $p$. In a doubly-linked skip list, each node stores a predecessor pointer and a successor pointer for each list in which it appears, for an average of $\frac{2}{1-p}$ pointers per node. The lists at the higher level act as “express lanes” that allow the sequence of nodes to be traversed quickly. Searching for a node with a particular key involves searching first in the highest level, and repeatedly dropping down a level whenever it becomes clear that the node is not in the current level. Considering the search path in reverse shows that no more than $\frac{1}{1-p}$ nodes are searched on average per level, giving an average search time of $O\left(\log n \frac{1}{(1-p)\log \frac{1}{p}}\right)$ with $n$ nodes at level 0. Skip lists have been extensively studied [Pugh 1990; Papadakis et al. 1990; Devroye 1992; Kirschenhofer and Prodinger 1994; Kirschenhofer et al. 1995], and because they require no global balancing operations are particularly useful in parallel systems [Gabarró et al. 1996; Gabarró and Meseguer 1997].

![Fig. 1. A skip list with $n = 6$ nodes and $\lceil \log n \rceil = 3$ levels.](image)
Skip lists

• Build several lists at different “skip” steps

• $O(n)$ list
• Level 1: $\sim n/2$
• Level 2: $\sim n/4$
• ...
• Level $\log n$ $\sim 2$-3 elements...
Skip List

typedef struct nodeStructure *node;
typedef struct nodeStructure{
    keyType key;
    valueType value;
    node forward[1]; /* variable sized array of forward pointers */
};
Skip Lists

$S_3\vdash -\infty \rightarrow +\infty$

$S_2\vdash -\infty \rightarrow 15 \rightarrow +\infty$

$S_1\vdash -\infty \rightarrow 15 \rightarrow 23 \rightarrow +\infty$

$S_0\vdash -\infty \rightarrow 10 \rightarrow 15 \rightarrow 23 \rightarrow 36 \rightarrow +\infty$
Outline and Reading

• What is a skip list (§3.5)
• Operations
  – Search (§3.5.1)
  – Insertion (§3.5.2)
  – Deletion (§3.5.2)
• Implementation
• Analysis (§3.5.3)
  – Space usage
  – Search and update times
What is a Skip List

• A skip list for a set $S$ of distinct (key, element) items is a series of lists $S_0, S_1, \ldots, S_h$ such that
  – Each list $S_i$ contains the special keys $+\infty$ and $-\infty$
  – List $S_0$ contains the keys of $S$ in nondecreasing order
  – Each list is a subsequence of the previous one, i.e.,
    \[ S_0 \subseteq S_1 \subseteq \ldots \subseteq S_h \]
  – List $S_h$ contains only the two special keys

• We show how to use a skip list to implement the dictionary ADT
• We search for a key $x$ in a skip list as follows:
  – We start at the first position of the top list
  – At the current position $p$, we compare $x$ with $y \leftarrow key(after(p))$
    \begin{itemize}
    \item $x = y$: we return $element(after(p))$
    \item $x > y$: we “scan forward”
    \item $x < y$: we “drop down”
    \end{itemize}
  – If we try to drop down past the bottom list, we return $NO\_SUCH\_KEY$
• Example: search for 78
Randomized Algorithms

• A randomized algorithm performs coin tosses (i.e., uses random bits) to control its execution
• It contains statements of the type
  \[ b \leftarrow \text{random()} \]
  \[
  \text{if } b = 0 \\
  \text{do A ...} \\
  \text{else } \{ b = 1 \} \\
  \text{do B ...}
  \]
• Its running time depends on the outcomes of the coin tosses
• We analyze the expected running time of a randomized algorithm under the following assumptions
  – the coins are unbiased, and
  – the coin tosses are independent
• The worst-case running time of a randomized algorithm is often large but has very low probability (e.g., it occurs when all the coin tosses give “heads”)
• We use a randomized algorithm to insert items into a skip list
• To insert an item \((x, o)\) into a skip list, we use a randomized algorithm:
  – We repeatedly toss a coin until we get tails, and we denote with \(i\) the number of times the coin came up heads
  – If \(i \geq h\), we add to the skip list new lists \(S_{i+1}, \ldots, S_{i+h}\), each containing only the two special keys
  – We search for \(x\) in the skip list and find the positions \(p_0, p_1, \ldots, p_i\) of the items with largest key less than \(x\) in each list \(S_0, S_1, \ldots, S_i\)
  – For \(j \leftarrow 0, \ldots, i\), we insert item \((x, o)\) into list \(S_j\) after position \(p_j\)

• Example: insert key 15, with \(i = 2\)
Deletion

• To remove an item with key $x$ from a skip list, we proceed as follows:
  – We search for $x$ in the skip list and find the positions $p_0, p_1, \ldots, p_i$ of the items with key $x$, where position $p_j$ is in list $S_j$
  – We remove positions $p_0, p_1, \ldots, p_i$ from the lists $S_0, S_1, \ldots, S_i$
  – We remove all but one list containing only the two special keys

• Example: remove key 34
Implementation

• We can implement a skip list with quad-nodes
• A quad-node stores:
  – item
  – link to the node before
  – link to the node after
  – link to the node below
  – link to the node after
• Also, we define special keys PLUS_INF and MINUS_INF, and we modify the key comparator to handle them
Space Usage

• The space used by a skip list depends on the random bits used by each invocation of the insertion algorithm

• We use the following two basic probabilistic facts:
  
  **Fact 1:** The probability of getting \( i \) consecutive heads when flipping a coin is \( 1/2^i \)

  **Fact 2:** If each of \( n \) items is present in a set with probability \( p \), the expected size of the set is \( np \)

• Consider a skip list with \( n \) items
  
  – By Fact 1, we insert an item in list \( S_i \) with probability \( 1/2^i \)
  
  – By Fact 2, the expected size of list \( S_i \) is \( n/2^i \)

• The expected number of nodes used by the skip list is

\[
\sum_{i=0}^{h} \frac{n}{2^i} = n \sum_{i=0}^{h} \frac{1}{2^i} < 2n
\]

Thus, the expected space usage of a skip list with \( n \) items is \( O(n) \)
• The running time of the search an insertion algorithms is affected by the height $h$ of the skip list.

• We show that with high probability, a skip list with $n$ items has height $O(\log n)$.

• We use the following additional probabilistic fact:

  Fact 3: If each of $n$ events has probability $p$, the probability that at least one event occurs is at most $np$.

• Consider a skip list with $n$ items:
  – By Fact 1, we insert an item in list $S_i$ with probability $1/2^i$.
  – By Fact 3, the probability that list $S_i$ has at least one item is at most $n/2^i$.

• By picking $i = 3\log n$, we have that the probability that $S_{3\log n}$ has at least one item is at most $n/2^{3\log n} = n/n^3 = 1/n^2$.

• Thus a skip list with $n$ items has height at most $3\log n$ with probability at least $1 - 1/n^2$. 

Height
Search and Update Times

• The search time in a skip list is proportional to
  – the number of drop-down steps, plus
  – the number of scan-forward steps
• The drop-down steps are bounded by the height of the skip list and thus are $O(\log n)$ with high probability
• To analyze the scan-forward steps, we use yet another probabilistic fact:
  **Fact 4:** The expected number of coin tosses required in order to get tails is 2
• When we scan forward in a list, the destination key does not belong to a higher list
  – A scan-forward step is associated with a former coin toss that gave tails
• By Fact 4, in each list the expected number of scan-forward steps is 2
• Thus, the expected number of scan-forward steps is $O(\log n)$
• We conclude that a search in a skip list takes $O(\log n)$ expected time
• The analysis of insertion and deletion gives similar results
Summary

- A skip list is a data structure for dictionaries that uses a randomized insertion algorithm.

- In a skip list with \( n \) items:
  - The expected space used is \( O(n) \).
  - The expected search, insertion, and deletion time is \( O(\log n) \).

- Using a more complex probabilistic analysis, one can show that these performance bounds also hold with high probability.

- Skip lists are fast and simple to implement in practice.
Skip graphs


Abstract

Skip graphs are a novel distributed data structure, based on skip lists, that provide the full functionality of a balanced tree in a distributed system where resources are stored in separate nodes that may fail at any time. They are designed for use in searching peer-to-peer systems, and by providing the ability to perform queries based on key ordering, they improve on existing search tools that provide only hash table functionality. Unlike skip lists or other tree data structures, skip graphs are highly resilient, tolerating a large fraction of failed nodes without losing connectivity. In addition, constructing, inserting new nodes into, searching a skip graph, and detecting and repairing errors in the data structure introduced by node failures can be done using simple and straightforward algorithms.

- SODA 2003 proceedings version: PS, PDF.
- Journal version: PDF.

BibTeX

@article{Aspnes2007,
  title="Skip graphs",
  author="James Aspnes and Gauri Shah",
  journal="ACM Transactions on Algorithms",
  month=nov,
  year=2007,
  volume=3,
  number=4,
  pages={37}
}

Consolidated BibTeX file

Return to James Aspnes's publications
Return to James Aspnes's home page
Conclusions

• Abstract data types **hide implementations**
• Important is the functionality of the ADT
• *Data structures* and *algorithms* determine the speed of the operations on data
• Linear data structures provide good versatility
• Sorting – a most typical need/algorithm
• Sorting in $O(n \log n)$  Merge Sort, Quicksort
• Solving Recurrences – means to analyse
• Skip lists – $\log n$ **randomised** data structure