Big-Oh notation classes

<table>
<thead>
<tr>
<th>Class</th>
<th>Informal</th>
<th>Intuition</th>
<th>Analogy</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f(n) \in O(g(n)) )</td>
<td>( f ) is dominated by ( g )</td>
<td>Strictly below ( &lt; )</td>
<td></td>
</tr>
<tr>
<td>( f(n) \in \Omega(g(n)) )</td>
<td>Bounded from above</td>
<td>Upper bound ( \leq )</td>
<td></td>
</tr>
<tr>
<td>( f(n) \in \Theta(g(n)) )</td>
<td>Bounded from above and below</td>
<td>“equal to” ( = )</td>
<td></td>
</tr>
<tr>
<td>( f(n) \in \omega(g(n)) )</td>
<td>Bounded from below</td>
<td>Lower bound ( \geq )</td>
<td></td>
</tr>
</tbody>
</table>

Linear, sequential, ordered, list ...

Physical ordered list \(~\) array

- Memory /address/
  - Garbage collection

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

- Disk
  - Disk fragmentation

Data structures


- [https://en.wikipedia.org/wiki/List_of_terms RELATING to algorithms and data structures](https://en.wikipedia.org/wiki/List_of_terms RELATING to algorithms and data structures)

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

Memory, disk, tape etc – is an ordered sequentially addressed media.
Linear data structures: Lists

- Doubly linked list
- Array list
- Linked list
- Self-organizing list
- Skip list
- Unrolled linked list
- VList

Lists: Array

- Xor linked list
- Zipper
- Doubly connected edge list
- Difference list

Multiple lists, 2-D-arrays, etc...

2D array

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

Primitive & composite types

<table>
<thead>
<tr>
<th>Primitive types</th>
<th>Composite types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boolean</td>
<td>Array</td>
</tr>
<tr>
<td>Char</td>
<td>Record (also called tuple or struct)</td>
</tr>
<tr>
<td>Int</td>
<td>Union</td>
</tr>
<tr>
<td>Float</td>
<td>Tagged union (also called a variant, variant record, discriminated union, or disjoint union)</td>
</tr>
<tr>
<td>Double</td>
<td>Plain old data structure</td>
</tr>
<tr>
<td>String</td>
<td></td>
</tr>
<tr>
<td>Enumerated type</td>
<td></td>
</tr>
</tbody>
</table>

Abstract data types:

- Dictionary
- Stack (LIFO)
- Queue (FIFO)
- Queue (double-ended)
- Priority queue (fetch highest-priority object)
- ...

Abstract Data Types (ADT)

- Some common ADTs, which have proved useful in a great variety of applications, are

<table>
<thead>
<tr>
<th>Container</th>
<th>Stack</th>
</tr>
</thead>
<tbody>
<tr>
<td>List</td>
<td>Graph</td>
</tr>
<tr>
<td>Set</td>
<td>Queue</td>
</tr>
<tr>
<td>Multiset</td>
<td>Priority queue</td>
</tr>
<tr>
<td>Map</td>
<td>Double-ended queue</td>
</tr>
<tr>
<td>Multimap</td>
<td>Double-ended priority queue</td>
</tr>
</tbody>
</table>

- Each of these ADTs may be defined in many ways and variants, not necessarily equivalent.

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
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
  - Extendable Hashing

Linear data structures

- **Arrays**
  - Hashed array
  - Heightmap
  - Lookup table
  - Matrix
  - Parallel array
  - Sorted array
  - Sparse array
  - Sparse matrix
  - Suffix tree
  - Variable-length array

- **Lists**
  - Doubly linked list
  - Linked list
  - Self-organizing list
  - Skip list
  - Unrolled linked list
  - VIList
  - XOR linked list
  - Zipper
  - Doubly connected edge list

Trees...

- **Hashe**
  - Bloom filter
  - Distributed hash table
  - Hash array mapped trie
  - Hash list
  - Hash table
  - Hash trie
  - Koode
  - Prefix hash tree

Hashes, Graphs, Other

- **Hashes**
  - Adjacency list
  - Adjacency matrix
  - Graph-structured stack
  - Scene graph
  - Binary decision diagram
  - Zero suppressed decision diagram
  - AND-inverter graph
  - Directed graph
  - Directed acyclic graph

- **Graphs**
  - Propositional directed acyclic graph
  - Multigraph
  - Hypergraph

- **Other**
  - Lightmap
  - Winged edge
  - Quad-edge
  - Routing table
  - Symbol table

Lists: Array

```
0 1 size MAX_SIZE
3 6 7 5 2
```

```
0 1 size
3 6 7 5 2
```

```
0 1 size
3 6 7 5 2
```

```
0 1 size
3 6 7 5 2
```

Lists: Array

```
0 1 size
3 6 7 8 5 2
```

```
0 1 size
3 6 7 8 5 2
```

```
0 1 size
3 6 7 8 5 2
```

```
0 1 size
3 6 7 8 5 2
```

```
0 1 size
3 6 7 8 5 2
```

```
0 1 size
3 6 7 8 5 2
```
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

Linked lists

- Array indexed from 0 to n - 1:
  - \( k = 1 \), \( 1 \leq k < n \), \( k = n \)
  - access/change the \( k \)th element \( O(1) \), \( O(n) \), \( O(1) \)
  - insert before or after the \( k \)th element \( O(n) \), \( O(n) \), \( O(1) \)
  - delete the \( k \)th element \( O(n) \), \( O(n) \), \( O(1) \)

- Singly-linked list with head and tail pointers
  - \( k = 1 \), \( 1 \leq k < n \), \( k = n \)
  - access/change the \( k \)th element \( O(1) \), \( O(n) \), \( O(1) \)
  - insert before or after the \( k \)th element \( O(n) \), \( O(n) \), \( O(1) \)
  - delete the \( k \)th element \( O(n) \), \( O(n) \), \( O(1) \)

- Doubly linked list
  - \( k = 1 \), \( 1 \leq k < n \), \( k = n \)
  - access/change the \( k \)th element \( O(1) \), \( O(n) \), \( O(1) \)
  - insert before or after the \( k \)th element \( O(n) \), \( O(n) \), \( O(1) \)
  - delete the \( k \)th element \( O(n) \), \( O(n) \), \( O(1) \)

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;
  ```

- Singly-linked

- Doubly linked

- Linked lists: add

- Linked lists: delete (+ garbage collection?)
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?
- node *p;
- p = new node;
- p->num = 5;
- p->word = "Ali";
- p->next = NULL;
- node *q;

Pointers in 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

Maintaining list of free objects

# 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

# Maintaining list of free objects
Multiple lists, single free list

head1=3 => 8, 4, 7
head2=6 => 3, 9
free =2

next  1  2  3  4  5  6  7
key   1  2  3  4  5  6  7
prev  2  7  8  9  4  5  6

XOR linked lists are a data structure used in computer programming. They take advantage of the bitwise exclusive disjunction (XOR) operation, here denoted by ⊕, 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.

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:

Example with 4 bits

A  1011
B  0110
C  1100

AC  0111  A xor C
ACC 1011  AC xor C == A

Truth table for all binary logical operators

Queue

- enqueue(x) - add to end
- dequeue() - fetch from beginning
- FIFO – First In First Out
- O(1) in all reasonable cases 😊

Queue (FIFO)

Queue (FIFO) (basic idea, does not contain all controls!)

Circular buffer

- 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

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==MAX_SIZE )
Empty : F<0 or F>=L

L F MAX_SIZE-1
3 6 7 5 2
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 * 2 - + => 5+((4*3)-2)
- Very simple to parse and interpret
- FORTH, Postscript are stack-based languages

Array based stack

- How to know how big a stack shall ever 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!

What about deletions?

- when n=32 -> 33 (copy 32, insert 1)
- delete: 33->32
  - should you delete (slots 33-64) immediately?

- 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**

```
Binary search – recursive
BinarySearch(A[0..N-1], value, low, high)
{
    if (high < low)
        return -1 // not found
    mid = (low + high) / 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

• \( x \leftrightarrow A[18] \) ? <
• \( x \leftrightarrow A[9] \) ? >
• \( x \leftrightarrow A[13] \) ? ==

• \( O(\lg n) \)

Sorting

• given a list, arrange values so that \( L[1] \leq L[2] \leq \ldots \leq L[n] \)
• \( n \) 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...

Decision-tree example

Sort \( (a_1, a_2, a_3) \)

\[ = (9, 4, 6) : \]

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.

Lower bound for decision-tree sorting

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

**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, \( n! \leq 2^h \).

\[ \therefore h \geq \lg(n!) \quad (\lg \text{ is mono. increasing}) \]

\[ \geq \lg \left(\left(\frac{n}{e}\right)^n\right) \quad \text{(Stirling’s formula)} \]

\[ = n \lg n - n \lg e \]

\[ = \Omega(n \lg n) \]

Decision tree model

\[ \text{Proof: } \log(n!) = \Omega(n \log n) \]

• \( \log(n!) = \log n + \log(n-1) + \log(n-2) + \cdots + \log(1) \)

\[ \geq n/2 \cdot \log(n/2) \]

\[ = \Omega(n \log n) \]

Half of elements are larger than \( \log(n/2) \)
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.

\[
\left\lceil \log_2 n! \right\rceil \geq \log_2 n! \\
\geq \sum_{i=1}^{n} \log_2 i \\
\geq \sum_{i=1}^{n/2} \log_2 n/2 \\
\geq n/2 \log_2 n/2 \\
= \Omega(n \log n).
\]

Merge sort

**Merge-Sort(A,p,r)**

if \( p < r \)

then \( q = (p+r)/2 \) \hspace{1cm} // floor

\[
\begin{align*}
\text{Merge-Sort}(A, p, q) \\
\text{Merge-Sort}(A, q+1, r) \\
\text{Merge}(A, p, q, r)
\end{align*}
\]

It was invented by John von Neumann in 1945.

Example

- Applying the merge sort algorithm:

  ![Merge sort example](image)

  Wikipedia / viz.

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.
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.

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/

Quicksort

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

Divide and conquer

Quick sort an \( n \)-element array:

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

2. Conquer: Recursively sort the two subarrays.


Key: Linear-time partitioning subroutine.
Pseudocode for quicksort

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

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

Partitioning subroutine

\[
\text{PARTITION}(A, p, q) \\
x \leftarrow A[p] \\
i \leftarrow p \\
\text{for } j \leftarrow p + 1 \text{ to } q \\
\text{do if } A[j] \leq x \\
\text{then } i \leftarrow i + 1 \\
\text{exchange } A[i] \leftrightarrow A[j] \\
\text{return } i
\]

Invariant: \begin{tabular}{c|c|c|c|c|c}
\hline
x & \leq x & \geq x & ? \\
\hline
p & i & j & q \\
\hline
\end{tabular}

Running time = \(O(n)\) for \(n\) elements.

Partitioning version 2

\[
\text{pivot} = A[R]; \\
i = L; j = R-1; \\
\text{while } (i < j) \\
\text{while } (A[i] < \text{pivot}) i++ ; \text{// will stop at pivot latest} \\
\text{while } (i < j \text{ and } A[j] \geq \text{pivot}) j--; \\
\text{if } (i < j) \{ \text{swap } A[i], A[j] ; i++; j-- \} \\
A[R] = A[i]; \\
A[i] = \text{pivot}; \\
\text{return } i;
\]

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

\[(\text{For intuition only!})\]

If we’re lucky, \text{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 \(1 : 9\)?

\[
T(n) = T(\lfloor n/10 \rfloor) + T(\lceil n/10 \rceil) + \Theta(n)
\]

What is the solution to this recurrence?
Choice of pivot in Quicksort

- 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.
• 2-pivot version of Quicksort
  – split in 3 regions!
• Handle equal values (equal to pivots) as „third region“

---

**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.

---

**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 \( \text{PARTITION} \) generates a \( k:n-k-1 \) 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.
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] 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] E[T(k) + T(n-k-1) + \Theta(n)]

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

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

Hairy recurrence

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

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

\[ E[T(n)] \leq \sum_{k=2}^{\log_2 n} a k \log k + \Theta(n) \]
\[ = \frac{2a}{n} \left( \frac{n^2 \log n - \frac{1}{2} n^2}{8} \right) + \Theta(n) \]
\[ = an \log n - \frac{an}{4} - \Theta(n) \]
\[ \leq an \log n, \]
if \( a \) is chosen large enough so that \( an/4 \) dominates the \( \Theta(n) \).

Alternative materials

- Quicksort average case analysis
  - [http://eid.ee/10z](http://eid.ee/10z)
  - [https://secweb.cs.odu.edu/~zeil/cs361/web/website/lectures/math/pages/qsort05.html](https://secweb.cs.odu.edu/~zeil/cs361/web/website/lectures/math/pages/qsort05.html)
- [http://eid.ee/10y](http://eid.ee/10y) - MIT Open Courseware
  - Asymptotic notation, Recurrences,
    - Substitution Master Method

The master method

The master method applies to recurrences of the form

\[ T(n) = a T(n/b) + f(n), \]
where \( a \geq 1, b > 1 \), and \( f \) is asymptotically positive.

\[ n \quad \text{vs} \quad \log_a n \]

\[ n^a \quad \text{vs} \quad f(n) \]
Three common cases

Compare \( f(n) \) with \( n^{\log_b a} \):

1. \( f(n) = O(n^{\log_b a - \epsilon}) \) for some constant \( \epsilon > 0 \).
   - \( f(n) \) grows polynomially slower than \( n^{\log_b a} \) (by an \( n^{\epsilon} \) factor).
   
   **Solution:** \( T(n) = \Theta(n^{\log_b a}) \).

Three common cases (cont.)

Compare \( f(n) \) with \( n^{\log_b a} \):

2. \( f(n) = \Omega(n^{\log_b a + \epsilon}) \) for some constant \( \epsilon > 0 \).
   - \( f(n) \) grows polynomially faster than \( n^{\log_b a} \) (by an \( n^{\epsilon} \) factor),
   
   and \( f(n) \) satisfies the regularity condition that
   \( af(n/b) \leq cf(n) \) for some constant \( c < 1 \).
   
   **Solution:** \( T(n) = \Theta(f(n)) \).

Examples

**Example 4.1**

\[ T(n) = 4T(n/2) + n \]
\[ a = 4, \ b = 2 \Rightarrow n^{\log_b a} = n^2; f(n) = n. \]

**Case 1:** \( f(n) = O(n^{2-\epsilon}) \) for \( \epsilon = 1 \).

\[ T(n) = \Theta(n^2). \]

**Example 4.2**

\[ T(n) = 4T(n/2) + n^3 \]
\[ a = 4, \ b = 2 \Rightarrow n^{\log_b a} = n^2; f(n) = n^3. \]

**Case 3:** \( f(n) = \Omega(n^{2+\epsilon}) \) for \( \epsilon = 1 \)

and \( 4(n/2)^3 \leq cn^3 \) (reg. cond.) for \( c = 1/2 \).

\[ T(n) = \Theta(n^3). \]

Examples

**Example 5.1**

\[ T(n) = 4T(n/2) + n^3 \]
\[ a = 4, \ b = 2 \Rightarrow n^{\log_b a} = n^2; f(n) = n^3. \]

**Case 3:** \( f(n) = \Omega(n^{2+\epsilon}) \) for \( \epsilon = 1 \)

and \( 4(n/2)^3 \leq cn^3 \) (reg. cond.) for \( c = 1/2 \).

\[ T(n) = \Theta(n^3). \]

**Example 5.2**

\[ T(n) = 4T(n/2) + n^2 \log n \]
\[ a = 4, \ b = 2 \Rightarrow n^{\log_b a} = n^2; f(n) = n^2 \log n. \]

Master method does not apply. In particular, for every constant \( \epsilon > 0 \), we have \( n^\epsilon = o(\log n) \).
**Idea of master theorem**

**Recursion tree:**

- $f(n)$
- $f(n/b)$
- $f(n/b)$
- $\cdots$
- $f(n/b)$
- $f(n/b^2)$
- $f(n/b^2)$
- $\cdots$
- $f(n/b^2)$
- $T(1)$

$h = \log_b n$

#leaves $= a^h$

- $= a^{\log_b n}$
- $= n^{\log_b a}$

CASE 1: The weight increases geometrically from the root to the leaves. The leaves hold a constant fraction of the total weight.

$\Theta(n^{\log_b a})$

CASE 2: $(k = 0)$ The weight is approximately the same on each of the $\log_b n$ levels.

$\Theta(n^{\log_b a} \log n)$

CASE 3: The weight decreases geometrically from the root to the leaves. The root holds a constant fraction of the total weight.

$\Theta(f(n))$
Back to sorting

We can sort in $O(n \log n)$

- using comparisons ($<$ $<=$ $=>$ $>$) we can not do better than $O(n \log n)$
- Is that the best we can do?

How fast can we sort n integers?

- Sort people by their sex? (F/M, 0/1)
- Sort people by year of birth?

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 ← 1 to k
    do C[i] ← 0
for j ← 1 to n
    do C[A[j]] + 1 → C[i] = |{key = i}|
for i ← 2 to k
    do C[i] ← C[i] + C[i-1]
    → C[i] = |{key ≤ i}|
for j ← n dowato 1
    do B[C[A[j]]] ← A[j]
    C[A[j]] ← C[A[j]] - 1
```

Loop 1

```
A:  4 1 3 4 3
C:  0 0 0 0

for i ← 1 to k
    do C[i] ← 0
```

B:  

```
A:  4 1 3 4 3
C:  0 0 0 0

for i ← 1 to k
    do C[i] ← 0
```
**Loop 2**

A: 4 1 3 4 3
B: 

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

**Loop 3**

A: 4 1 3 4 3
B: 

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

**Loop 4**

A: 4 1 3 4 3
B: 3 4

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

**Analysis**

\[
\begin{align*}
\Theta(k) & \quad \text{for } i \leftarrow 1 \text{ to } k \\
& \text{do } C[i] \leftarrow 0 \\
\Theta(n) & \quad \text{for } j \leftarrow 1 \text{ to } n \\
& \text{do } C[A[j]] \leftarrow C[A[j]] + 1 \\
\Theta(k) & \quad \text{for } i \leftarrow 2 \text{ to } k \\
& \text{do } C[i] \leftarrow C[i] + C[i-1] \\
\Theta(n) & \quad \text{for } j \leftarrow n \text{ downto 1} \\
& \text{do } B[C[A[j]]] \leftarrow A[j] \\
& \text{do } C[A[j]] \leftarrow C[A[j]] - 1 \\
\Theta(n + k) & \\
\end{align*}
\]

**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.

A: 4 1 3 4 3
B: 1 3 3 4 4

**Exercise:** What other sorts have this property?
Radix sort

- **Origin**: Herman Hollerith's card-sorting machine for the 1890 U.S. Census. (See Appendix A.)
- 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 (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

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>329</td>
<td>720</td>
<td>720</td>
<td>329</td>
</tr>
<tr>
<td>457</td>
<td>355</td>
<td>329</td>
<td>355</td>
</tr>
<tr>
<td>657</td>
<td>436</td>
<td>436</td>
<td>436</td>
</tr>
<tr>
<td>839</td>
<td>457</td>
<td>839</td>
<td>457</td>
</tr>
<tr>
<td>436</td>
<td>657</td>
<td>355</td>
<td>657</td>
</tr>
<tr>
<td>720</td>
<td>329</td>
<td>457</td>
<td>720</td>
</tr>
<tr>
<td>355</td>
<td>839</td>
<td>657</td>
<td>839</td>
</tr>
</tbody>
</table>

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

| 8 | 8 | 8 | 8 |

\( r = 8 \Rightarrow b/r = 4 \) passes of counting sort on base-2\(^4\) 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

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

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

- Increasing $r$ means fewer passes, but as $r \gg \lg n$, 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 $\Theta(dn)$ 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)
Why not from left to right?

- Swap '0' with first '1'
- Idea 1: recursively sort first and second half
  - Exercise?

Bitwise sort left to right

- Idea 2:
  - 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

```c
void bitwise_sort(SORTTYPE *ARRAY, int size)
{
    int i, j, tmp, curbit;
    register SORTTYPE mask, curbit, group;
    nrbits = sizeof(SORTTYPE) * 8;
    curbit = 1 << (nrbits - 1); /* set most significant bit 1 */
    mask = 0;
    do { /* For each bit */
        i = 0;
        new_mask:
        for( ; (i < size) && (ARRAY[i] & curbit) == 0; 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) == 0 ) {    /* 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... */
}
```

Bitwise from left to right

- 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

- The sort input records must be 100 bytes in length, with the first 10 bytes being a random key
- Minutecore – 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

http://sortbenchmark.org/
Sort Benchmark

- [http://sortbenchmark.org/](http://sortbenchmark.org/)
- Sort Benchmark Home Page
- We have a new benchmark called newGraySort, 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 (uninterrupted) 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[(n+1)/2] or A[(n+1)/2]

Q: Find i’th value in unsorted data

A. O(n) 
B. O(n log log n) 
C. O(n log n) 
D. O(n log^2 n) 

Minimum

```plaintext
Minimum(A)  
1 min = A[1]  
2 for i = 2 to length(A)  
3 if min > A[i]  
4 then min = A[i]  
5 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

\[ 3 \left\lfloor \frac{n}{2} \right\rfloor \]

Selection in expected O(n)

Randomised-select( \( A, p, r, i \) )
if \( p=r \) then return \( A[p] \)
\( q = \) Randomised-Partition\( (A,p,r) \)
\( k = q - p + 1 \) \hspace{1cm} // 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 in unsorted – expected \( O(n) \)

- Find \( i' \) th value: worst case \( O(n) \) – see CLRS

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

Q: search for a value X in linked list?

A. \( O(1) \)
B. \( O(\log n) \)
C. \( O(n) \)

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 Lists

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., $S_0 \supseteq S_1 \supseteq \cdots \supseteq S_h$
- List $S_h$ contains only the two special keys

We show how to use a skip list to implement the dictionary ADT

typedef struct nodeStructure *node;

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

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., $S_0 \supseteq S_1 \supseteq \cdots \supseteq S_h$
  - List $S_h$ contains only the two special keys

We show how to use a skip list to implement the dictionary ADT

Skip List

typedef struct nodeStructure *node;

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

Illustration of lists

Key
Right[ ] – right links in array
Left[ ] – left links in array, array size – how high is the list

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...
Search

- We search for a key \( x \) in a skip list as follows:
  - We start at the first position of the top list
  - If the current position \( p \), we compare \( x \) with \( y = \text{key}(p) \)
    - If \( x > y \), we return \( \text{element}(p) \)
    - If \( x < y \), we "drop down"
  - If we try to drop down past the bottom list, we return \( \text{NO_SUCH_KEY} \)
- 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
    - \( b \leftarrow \text{random()} \)
    - \( \text{if } b = 0 \text{ then } \ldots \text{ else if } b = 1 \text{ then } \ldots \)
    - \( \text{do } A \ldots \text{ else } D \ldots \)
  - 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

Insertion

- To insert an item \((x, e)\) 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 \( i \geq \theta \), we add to the skip list new lists \( S_1, S_2, \ldots, S_{i-2}, S_{i-1} \), each containing only the two special keys
  - We search for \( x \) in the skip list and find the positions \( p_1, p_2, \ldots, p_k \) of the items with largest key less than \( x \) in each list \( S_1, S_2, \ldots, S_{i-1} \)
  - For \( j = 0, \ldots, i \), we insert item \((x, e)\) 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 \( p_1, p_2, \ldots, p_k \) of the items with key \( x \), where position \( p_0 \) is in list \( S_n \)
  - We remove positions \( p_0, p_1, \ldots, p_k \) from the lists \( S_n, S_{n-1}, \ldots, S_1 \)
  - We remove all but one list containing only the two special keys
- Example: remove key 34

Implementation v2

- We can implement a skip list with quad nodes
- A quad node stores:
  - Item
  - Link to the node before
  - Link to the node below
  - Link to the node after
- Also, we define special keys \( \text{PLUS_INF} \) and \( \text{MINUS_INF} \), 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 \( \alpha \) items is present in a set with probability \( p \), the expected size of the set is \( \alpha p \)
  - We consider a skip list with \( \alpha \) items
    - By Fact 1, we insert an item in list \( S_n \) with probability \( 1/2^{\theta} \)
    - By Fact 2, the expected size of list \( S_n \) is \( \alpha/2^\theta \)
  - The expected number of nodes used by the skip list is
    \[
    \sum_{k=0}^{\theta} \frac{\alpha}{2^k} + \sum_{k=\theta}^{\infty} \frac{\alpha}{2^k} < 2\alpha
    \]
  - Thus, the expected space usage of a skip list with \( \alpha \) items is \( O(\alpha) \)
Height

• The running time of the search and 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 $1 - (1 - p)^n$.

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 $2^{i-1}/2^i = 1/2^{i-1}$.
- Thus, a skip list with $n$ items has height at most $\log_2 n$.

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

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 $O(\log n)$.
- 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.

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.