

**Facts**

$$\sum_{k=0}^n ar^k = \frac{ar^{n+1} - a}{r - 1} \quad \text{for } r \notin \{0, 1\}$$

$\mathcal{O}(1) < \mathcal{O}((\log n)^c) < \mathcal{O}(\log n)$   
 $< \mathcal{O}(\log^2 n) < \mathcal{O}(n) < \mathcal{O}(n \log n)$   
 $< \mathcal{O}(n^c) < \mathcal{O}(n^{\log n}) < \mathcal{O}(n!) < \mathcal{O}(n^n)$

**Loop Invariant**

A property that holds before and after each loop iteration.

**Initialization:** Holds before the first iteration.

**Maintenance:** If it holds before an iteration, it holds after.

**Termination:** When the loop ends, the invariant helps prove correctness.

**Divide and Conquer**

An algorithmic paradigm with three steps:

1. **Divide:** Split the problem into smaller subproblems.

2. **Conquer:** Solve each subproblem recursively.

3. **Combine:** Merge the subproblem solutions into the final result. The recurrence relation is:

$$T(n) = \begin{cases} \Theta(1) & \text{if } n \leq c, \\ aT(n/b) + D(n) + C(n) & \text{otherwise,} \end{cases}$$

$a$ : number of subproblems,  
 $n/b$ : size of each subproblem,  
 $D(n)$ : time to divide,  
 $C(n)$ : time to combine.

**Solving Recurrences**

To solve a recurrence using the substitution method:

1. **Guess** the solution's form (e.g.,  $\Theta(n^2)$ ).

2. **Prove** the upper bound by induction using a constant and the guessed form.

$$T(n) = 2T(n/2) + cn$$
$$\leq 2 \frac{an}{2} \log(n/e) + cn$$
$$= an \log n - an + cn$$
$$\leq an \log n$$
$$= \mathcal{O}(n \log n)$$

3. **Prove** the lower bound similarly.

4. **Conclude** that the guess is correct.

**Example:**

$$T(n) = T(n-1) + cn \Rightarrow \Theta(n^2)$$

**Master Theorem**

Let  $a \geq 1$ ,  $b > 1$ , and let  $T(n)$  be defined by the recurrence:

$$T(n) = aT(n/b) + f(n)$$

Then  $T(n)$  has the following asymptotic bounds:

1. If  $f(n) = \mathcal{O}(n^{\log_b a - \epsilon})$  for some  $\epsilon > 0$ ,

$$\text{then } T(n) = \Theta(n^{\log_b a}).$$

2. If  $f(n) = \Theta(n^{\log_b a} \log^k n)$  for some  $k \geq 0$ , then

$$T(n) = \Theta(n^{\log_b a} \log^{k+1} n).$$

3. If  $f(n) = \Omega(n^{\log_b a + \epsilon})$  for some  $\epsilon > 0$ ,

$$\text{and if } a f(n/b) \leq c f(n) \text{ for some } c < 1 \text{ and large } n,$$
$$\text{then } T(n) = \Theta(f(n)).$$

Special case:

$$T(n) = a_1 T(b_1 n) + a_2 T(b_2 n) + \dots + n^c$$

1. If  $a_1 b_1^c + a_2 b_2^c + \dots < 1$  then  $\Theta(n^c)$ .

2. If  $a_1 b_1^c + a_2 b_2^c + \dots = 1$  then  $\Theta(n^c \log n)$ .

3. If  $a_1 b_1^c + a_2 b_2^c + \dots > 1$  then  $\Theta(n^{\epsilon})$ , where  $a_1 b_1^{\epsilon} + a_2 b_2^{\epsilon} + \dots = 1$ .

**Injective Functions**

Let  $f: \{1, 2, \dots, q\} \rightarrow M$  be a function chosen uniformly at random, where  $|M| = m$ . If  $q > 1.78\sqrt{m}$ , then the probability that  $f$  is injective is at most  $\frac{1}{2}$ .

**Insertion Sort**

**Steps:**

1. Start with an empty (or trivially sorted) sublist.

2. Insert the next element in the correct position by comparing backwards.

3. Repeat for all elements.

**Complexity:**

Space:  $\Theta(n)$ , Time:  $\Theta(n^2)$

**Pseudocode:**

**Require:**  $A = \langle a_1, a_2, \dots, a_n \rangle$

1: for  $i \leftarrow 2$  to  $n$  do

2:    $key \leftarrow A[i]$

3:    $j \leftarrow i - 1$

4:   while  $j \geq 1$  and  $A[j] > key$  do

5:      $A[j+1] \leftarrow A[j]$

6:      $j \leftarrow j - 1$

7:   end while

8:    $A[j+1] \leftarrow key$

9: end for

**Heap Sort**

A heap is a nearly complete binary tree where each node satisfies the **max-heap property**. For every node  $i$ , its children have smaller or equal values.

The **height** of a heap is the length of the longest path from the root to a leaf.

**Useful Index Rules (array-based heap):**

• Root is at index  $A[1]$

• Left child of node  $i$ : index  $2i$

• Right child of node  $i$ : index  $2i + 1$

• Parent of node  $i$ : index  $\lfloor i/2 \rfloor$

**Complexity:**

Space:  $\Theta(n)$ , Time:  $\Theta(n \log n)$

**Pseudocode:**

1: procedure MAX-HEAPIFY( $A, i, n$ )

2:    $l \leftarrow \text{Left}(i)$

3:    $r \leftarrow \text{Right}(i)$

4:    $\text{largest} \leftarrow i$

5:   if  $l \leq n$  and  $A[l] > A[\text{largest}]$  then

6:      $\text{largest} \leftarrow l$

7:   end if

8:   if  $r \leq n$  and  $A[r] > A[\text{largest}]$  then

9:      $\text{largest} \leftarrow r$

10:   end if

11:   if  $\text{largest} \neq i$  then

12:     Exchange  $A[i] \leftrightarrow A[\text{largest}]$

13:     MAX-HEAPIFY( $A, \text{largest}, n$ )

14:   end if

15: end procedure

1: procedure BUILD-MAX-HEAP( $A[1, \dots, n]$ )

2:   BUILD-MAX-HEAP( $A$ )

3:   for  $i \leftarrow \lfloor n/2 \rfloor$  downto 1 do

4:     exchange  $A[i]$  with  $A[i]$

5:     MAX-HEAPIFY( $A, i, n$ )

6:   end for

7: end procedure

**Merge Sort**

Uses the **divide and conquer** paradigm.

**Complexity:**

Space:  $\Theta(n)$ , Time:  $\Theta(n \log n)$

**Pseudocode:**

1: procedure SORT( $A, p, r$ )

2:   if  $p < r$  then

3:      $q \leftarrow \lfloor (p+r)/2 \rfloor$

4:     SORT( $A, p, q$ )

5:     SORT( $A, q+1, r$ )

6:     MERGE( $A, p, q, r$ )

7:   end if

8: end procedure

1: procedure MERGE( $A, p, q, r$ )

2:    $n_1 \leftarrow q - p + 1$ ,  $n_2 \leftarrow r - q$

3:   Let  $L[1 \dots n_1 + 1]$ ,  $R[1 \dots n_2 + 1]$  be new arrays

4:   for  $i \leftarrow 1$  to  $n_1$  do  $L[i] \leftarrow A[p + i - 1]$

5:   end for

6:   for  $j \leftarrow 1$  to  $n_2$  do  $R[j] \leftarrow A[q + j]$

7:   end for

8:    $L[n_1 + 1]$ ,  $R[n_2 + 1] \leftarrow \infty$

9:    $i, j \leftarrow 1$

10:   for  $k \leftarrow p$  to  $r$  do

11:     if  $L[i] \leq R[j]$  then

12:        $A[k] \leftarrow L[i]$ ;  $i \leftarrow i + 1$

13:     else

14:        $A[k] \leftarrow R[j]$ ;  $j \leftarrow j + 1$

15:     end if

16:   end for

17: end procedure

**Stack**

A stack is a last-in/fist-out (LIFO) data-structure

**Supported operations:**

• **Push:** Insert an element at head.

• **Pop:** Retrieve head.  $\mathcal{O}(1)$

**Maximum Subarray Problem (Kadane's Algorithm)**

**Idea:** Iterate from left to right, maintaining: - **endingHereMax**: best subarray ending at current index - **currentMax**: best seen so far

**Observation:** At index  $j+1$ , the maximum subarray is either:

• the best subarray in  $A[1..j]$ , or

• a subarray ending at  $j+1$ , i.e.,  $A[i..j+1]$

**Formula:**

$$\text{maxSub}(A[1..n]) = \max(\text{maxSub}(A[1..n-1]), \text{max}_{i=1}^n \text{sum}(A[i..n]))$$

**Complexity:** Time  $\Theta(n)$ ,

Space  $\mathcal{O}(1)$

**Pseudocode:**

1: procedure SUBARRAY( $A[1..n]$ )

2:    $\text{current\_max} \leftarrow -\infty$

3:    $\text{ending\_here\_max} \leftarrow -\infty$

4:   for  $i \leftarrow 1$  to  $n$  do

5:      $\text{ending\_here\_max} \leftarrow \max(A[i], \text{ending\_here\_max} + A[i])$

6:      $\text{current\_max} \leftarrow \max(\text{current\_max}, \text{ending\_here\_max})$

7:   end for

8:   return  $\text{current\_max}$

9: end procedure

**Queue**

A queue is a first-in, first-out (FIFO) collection.

• **enqueue:** Insert an element at tail.

• **dequeue:** Retrieve head.  $\mathcal{O}(1)$

1: procedure ENQUEUE( $Q, x$ )

2:    $Q[Q.\text{tail}] \leftarrow x$

3:   if  $Q.\text{tail} = Q.\text{length}$  then

4:      $Q.\text{tail} \leftarrow 1$

5:   else

6:      $Q.\text{tail} \leftarrow Q.\text{tail} + 1$

7:   end if

8: end procedure

9: end procedure

10: procedure DEQUEUE( $Q$ )

11:    $x \leftarrow Q[Q.\text{head}]$

12:   if  $Q.\text{head} = Q.\text{length}$  then

13:      $Q.\text{head} \leftarrow 1$

14:   else

15:      $Q.\text{head} \leftarrow Q.\text{head} + 1$

16:   end if

17:   return  $x$

18: end procedure

**Dynamic Programming**

Two key approaches: **Top-down** and **Bottom-up**.

• **Top-down:** Starts from the problem  $n$  and solves subproblems recursively, storing results (memoization).

• **Bottom-up:** Starts from base cases (e.g., 0) and iteratively builds up to the final solution.

The core idea is to **remember previous computations** to avoid redundant work and save time.

**Hash Functions and Tables**

Tables are a special kind of collection that associate keys to values, allowing the following operations:

1. Insert a new key-value pair.

2. Delete a key-value pair.

3. Search for the value associated with a given key.

**Direct-Address Tables.** We define a function  $f: K \rightarrow \{1, \dots, |K|\}$  and create an array of size  $|K|$  where each position corresponds directly to a key, allowing constant-time access.

**Hash Tables.** Hash tables use space proportional to the number of stored keys  $|K'|$ , i.e.,  $\Theta(|K'|)$ , and support the above operations in expected time  $\mathcal{O}(1)$  in the average case. To achieve this, we define a hash function  $h: K \rightarrow \{1, \dots, M\}$  and use an array of size  $M$  where each entry contains a linked list of key-value pairs  $(k, v)$ .

**Strassen's Algorithm for Matrix Multiplication**

Instead of performing 8 recursive multiplications as in the naive divide-and-conquer matrix multiplication, Strassen's algorithm reduces it to 7, which improves the time complexity.

**Definitions:**

$$M_1 = (A_{11} + A_{22})(B_{11} + B_{22})$$

$$M_2 = (A_{21} + A_{22})B_{11}$$

$$M_3 = A_{11}(B_{12} - B_{22})$$

$$M_4 = A_{22}(B_{21} - B_{11})$$

$$M_5 = (A_{11} + A_{12})B_{22}$$

$$M_6 = (A_{21} - A_{11})(B_{11} + B_{12})$$

$$M_7 = (A_{12} - A_{22})(B_{21} + B_{22})$$

**Resulting matrix:**

$$C_{11} = M_1 + M_4 - M_5 + M_7$$

$$C_{12} = M_3 + M_5$$

$$C_{21} = M_2 + M_4$$

$$C_{22} = M_1 - M_2 + M_3 + M_6$$

**Complexity:**

$$\text{Time: } \Theta(n^{\log_2 7}) \approx \Theta(n^{2.81})$$

Space:  $\Theta(n^2)$

**Priority Queue**

A priority queue maintains a dynamic set  $S$  of elements, each with an associated key that defines its priority. At each operation, we can access the element with the highest key.

**Supported operations:**

• **Insertion:** Insert an element  $x$  into  $S$ .  $\mathcal{O}(\log n)$

• **Maximum:** Return the element in  $S$  with the largest key.  $\mathcal{O}(1)$

• **Extract-Max:** Remove and return the element with the largest key.  $\mathcal{O}(\log n)$

• **Increase-Key:** Increase the key of an element  $x$  to a new value  $k$  (assuming  $k \geq$  current key).  $\mathcal{O}(\log n)$

**Pseudocode:**

1: procedure HEAP-MAXIMUM( $S$ )

2:   return  $S[1]$

3: end procedure

1: procedure HEAP-EXTRACT-MAX( $S, n$ )

2:   if  $n < 1$  then

3:     error "heap underflow"

4:   end if

5:    $max \leftarrow S[1]$

6:    $S[1] \leftarrow S[n]$

7:    $n \leftarrow n - 1$

8:   MAX-HEAPIFY( $S, 1, n$ )

9:   return  $max$

10: end procedure

1: procedure HEAP-INCREASE-KEY( $S, i, key$ )

2:   if  $key < S[i]$  then

3:     error "new key is smaller than current key"

4:   end if

5:    $S[i] \leftarrow key$

6:   while  $i > 1$  and  $S[\text{Parent}(i)] < S[i]$  do

7:     exchange  $S[i]$  with  $S[\text{Parent}(i)]$

8:      $i \leftarrow \text{Parent}(i)$

9:   end while

10: end procedure

1: procedure MAX-HEAP-INSERT( $S, key, n$ )

2:    $n \leftarrow n + 1$

3:    $S[n] \leftarrow -\infty$

4:   HEAP-INCREASE-KEY( $S, n, key$ )

Longest Common Subsequence

Given as input two sequences  $X = \langle x_1, \dots, x_m \rangle$  and  $Y = \langle y_1, \dots, y_n \rangle$ , we want to find the longest common subsequence (not necessarily contiguous, but in order).

$$c[i, j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0, \\ c[i-1, j-1] + 1 & \text{if } x_i = y_j, \\ \max\{c[i-1, j], c[i, j-1]\} & \text{otherwise.} \end{cases}$$

```
1: procedure LCS-LENGTH(X, Y, m, n)
2:   let b[1...m][1...n] and c[0...m][0...n] be new tables
3:   for i ← 1 to m do
4:     c[i][0] ← 0
5:   end for
6:   for j ← 1 to n do
7:     c[0][j] ← 0
8:   end for
9:   for i ← 1 to m do
10:    for j ← 1 to n do
11:      if X[i] = Y[j] then
12:        c[i][j] ← c[i-1][j-1] + 1
13:      else
14:        b[i][j] ← "↖"
15:        if c[i-1][j] ≥ c[i][j-1] then
16:          c[i][j] ← c[i-1][j]
17:          b[i][j] ← "↑"
18:        else
19:          c[i][j] ← c[i][j-1]
20:          b[i][j] ← "←"
21:        end if
22:      end if
23:    end for
24:  end for
25:  return c and b
26: end procedure
```

Time complexity:  $\mathcal{O}(mn)$

Space complexity:  $\mathcal{O}(mn)$

Graph

A graph  $G = (V, E)$  consists of a vertex set  $V$  and an edge set  $E$  that contains (ordered) pairs of vertices.

Feature	Adjacency List	Adjacency Matrix
Time to list all vertices adjacent to $v$	$\mathcal{O}(\deg(v))$	$\mathcal{O}(V)$
Time to check if $(u, v) \in E$	$\mathcal{O}(V)$	$\mathcal{O}(1)$

Connectivity: A graph is said to be connected if every pair of vertices in the graph is connected.

Connected Component: A connected component is a maximal connected subgraph of an undirected graph.

Complete Graph: A complete graph is a simple undirected graph in which every pair of distinct vertices is connected by a unique edge.

Vertex Cut: A vertex cut or separating set of a connected graph  $G$  is a set of vertices whose removal renders  $G$  disconnected.

Breadth-First Search

Given as input a graph  $G = (V, E)$ , either directed or undirected, and a source vertex  $s \in V$ , we want to find  $v, d$ , the smallest number of edges (distance) from  $s$  to  $v$ , for all  $v \in V$ .

```
1: procedure BFS(V, E, s)
2:   for each  $v \in V \setminus \{s\}$  do
3:      $u, d \leftarrow \infty$ 
4:   end for
5:    $s, d \leftarrow 0$ 
6:   let  $Q$  be a new queue
7:   ENQUEUE( $Q, s$ )
8:   while  $Q \neq \emptyset$  do
9:      $u \leftarrow$  DEQUEUE( $Q$ )
10:    for each  $v \in G.\text{Adj}[u]$  do
11:      if  $v, d \geq \infty$  then
12:         $v, d \leftarrow u, d + 1$ 
13:        ENQUEUE( $Q, v$ )
14:      end if
15:    end for
16:  end while
17: end procedure
```

Time complexity:  $\mathcal{O}(|V| + |E|)$

Depth-First Search

Given as input, a graph  $G = (V, E)$ , either directed or undirected, we want to output two timestamps on each vertex:

- $v, d$  – discovery time (when  $v$  is first encountered),
- $v, f$  – finishing time (when all vertices reachable from  $v$  have been fully explored).

Each vertex has a color state:

- WHITE: undiscovered,
- GRAY: discovered but not finished,
- BLACK: fully explored.

```
1: procedure DFS(G)
2:   for each  $u \in G.V$  do
3:      $u.\text{color} \leftarrow$  WHITE
4:   end for
5:   time ← 0
6:   for each  $u \in G.V$  do
7:     if  $u.\text{color} =$  WHITE then
8:       DFS-VISIT( $G, u$ )
9:     end if
10:   end for
11: end procedure
```

```
12: procedure DFS-VISIT( $G, u$ )
13:   time ← time + 1
14:    $u, d \leftarrow$  time
15:    $u.\text{color} \leftarrow$  GRAY
16:   for each  $v \in G.\text{Adj}[u]$  do
17:     if  $v.\text{color} =$  WHITE then
18:       DFS-VISIT( $G, v$ )
19:     end if
20:   end for
21:    $u.\text{color} \leftarrow$  BLACK
22:   time ← time + 1
23:    $u, f \leftarrow$  time
24: end procedure
```

Time complexity:  $\mathcal{O}(|V| + |E|)$

Topological Sort

Given a directed acyclic graph (DAG)  $G = (V, E)$ , the goal is to produce a linear ordering of its vertices such that for every edge  $(u, v) \in E$ , vertex  $u$  appears before  $v$  in the ordering.

Key Properties:

- A graph is a DAG if and only if a DFS yields no back edges.
- The topological sort is obtained by performing DFS and ordering vertices in decreasing order of their finishing times.

Algorithm:

```
1. Run DFS on  $G$  to compute finishing times  $u, f$  for all  $v \in V$ .
2. Return the vertices sorted in descending order of  $u, f$ .
```

Running Time:  $\Theta(|V| + |E|)$ , same as DFS.

Strongly Connected Components

A strongly connected component (SCC) of a directed graph  $G = (V, E)$  is a maximal set of vertices  $C \subseteq V$  such that for every pair  $u, v \in C$ , there is a path from  $u$  to  $v$  and from  $v$  to  $u$ .

Transpose of a Graph: The transpose of  $G$ , denoted  $G^T = (V, E^T)$ , has all edges reversed:

$$E^T = \{(u, v) \mid (v, u) \in E\}$$

$G$  and  $G^T$  share the same SCCs. Computing  $G^T$  takes  $\Theta(|V| + |E|)$  time with adjacency lists.

Algorithm (Kosaraju's):

```
1. Run DFS on  $G$  to compute finishing times  $u, f$  for all  $u \in V$ .
2. Compute the transpose  $G^T$ .
3. Run DFS on  $G^T$ , but visit vertices in order of decreasing  $u, f$  (from step 1).
4. Each tree in the resulting DFS forest is one SCC.
```

Time Complexity:  $\Theta(|V| + |E|)$

Flow Network

We model the movement of flow through a network of edges, where each edge has a capacity – the maximum flow allowed. Our goal is to maximize the total flow from a source vertex  $s$  to a sink vertex  $t$ .

Flow Function: A flow is a function  $f: V \times V \rightarrow \mathbb{R}$  that satisfies:

```
1. Capacity Constraint: For all  $u, v \in V$ ,
0 ≤ f(u, v) ≤ c(u, v)
where c(u, v) is the capacity of edge (u, v).
2. Flow Conservation: For all  $u \in V \setminus \{s, t\}$ ,
∑_{v ∈ V} f(v, u) = ∑_{v ∈ V} f(u, v)
i.e., the total flow into u equals the total flow out of u (except for source and sink).
```

Value of the Flow:

$$|f| = \sum_{v \in V} f(s, v) - \sum_{v \in V} f(v, s)$$

This represents the total net flow out of the source  $s$ .

Ford-Fulkerson Method (1954)

The Ford-Fulkerson method finds the maximum flow from a source  $s$  to a sink  $t$  in a flow network  $G = (V, E)$ .

Algorithm:

```
1. Initialize flow f(u, v) = 0 for all (u, v) ∈ E.
2. While there exists an augmenting path p from s to t in the residual network G_f:
• Compute the bottleneck capacity c_f(p) (the minimum residual capacity along p).
• Augment flow f along p by c_f(p).
```

Residual Network: Given flow  $f$ , define residual capacity  $c_f$  as:

$$c_f(u, v) = \begin{cases} c(u, v) - f(u, v) & \text{if } (u, v) \in E \\ f(v, u) & \text{if } (v, u) \in E \\ 0 & \text{otherwise} \end{cases}$$

Then the residual graph is  $G_f = (V, E_f)$  where  $E_f = \{(u, v) \in V \times V : c_f(u, v) > 0\}$ .

Cuts and Optimality:

- A cut  $(S, T)$  of the network is a partition of  $V$  with  $s \in S, t \in T$ .
- The flow across the cut is:  $f(S, T) = \sum_{u \in S, v \in T} f(u, v) - \sum_{v \in T, u \in S} f(v, u)$
- The capacity of the cut is:  $c(S, T) = \sum_{u \in S, v \in T} c(u, v)$
- For any flow  $f$  and any cut  $(S, T)$ , we have:  $|f| \leq c(S, T)$ .

Max-Flow Min-Cut Theorem: The value of the maximum flow equals the capacity of the minimum cut.

Augmenting Path Is a path from the source to the sink in the residual graph such that every edge on the path has available capacity

Time complexity:  $\mathcal{O}(E|f|_{\max})$

Bipartite Graphs

A bipartite graph (or bigraph) is a graph  $G = (U, V, E)$  whose vertices can be partitioned into two disjoint sets  $U$  and  $V$  such that every edge connects a vertex from  $U$  to one in  $V$ .

- $U$  and  $V$  are called the parts of the graph.
- Bipartiteness can be tested using BFS:
  - Label the source  $s$  as even.
  - During BFS, label each unvisited neighbor of a vertex with the opposite parity (even  $\leftrightarrow$  odd).
  - If a conflict arises (a vertex is visited twice with the same parity), the graph is not bipartite.

Bipartite Match via Max-Flow

```
1. Add a source node s and connect it to all nodes in the left partition (say U) and do same for right part to sink.
2. For each edge (u, v) in the bipartite graph (with u ∈ U, v ∈ V), add a directed edge from u to v.
3. Assign a capacity of 1 to all edges.
4. Run Ford-Fulkerson from s to t.
```

Spanning Trees

Spanning Tree: A spanning tree  $T$  of an undirected graph  $G = (V, E)$  is a subgraph that:

- Includes all vertices of  $G$ .
- Is a tree: connected and acyclic.

Minimum Spanning Tree (MST): An MST is a spanning tree of a weighted graph with the minimum total edge weight among all spanning trees of the graph.

Key Properties:

- Every connected undirected graph has at least one MST.
- An MST connects all vertices using the lightest possible total edge weight without forming cycles.

Prim's Algorithm

Goal: Find a Minimum Spanning Tree (MST) of a connected, weighted undirected graph.

Idea:

```
1. Start from an arbitrary root vertex r.
2. Maintain a growing tree T, initialized with r.
3. Repeatedly add the minimum-weight edge that connects a vertex in T to a vertex outside T.
```

Data Structures: Uses a min-priority queue to select the next lightest edge crossing the cut.

```
1: procedure PRIM( $G, w, r$ )
2:   let  $Q$  be a new min-priority queue
3:   for each  $u \in G.V$  do
4:      $u.\text{key} \leftarrow \infty$ 
5:      $u.\pi \leftarrow$  NIL
6:   end for
7:   end for
8:   DECREASE-KEY( $Q, r, 0$ )
9:   while  $Q \neq \emptyset$  do
10:     $u \leftarrow$  EXTRACT-MIN( $Q$ )
11:    for each  $v \in G.\text{adj}[u]$  do
12:      if  $v \in Q$  and  $w(u, v) < v.\text{key}$  then
13:         $v.\pi \leftarrow u$ 
14:        DECREASE-KEY( $Q, v, w(u, v)$ )
15:      end if
16:    end for
17:  end while
18: end procedure
```

Runtime:  $\Theta(E \log V)$

$\Theta(E + V \log V)$  with Fibonacci heaps.

Kruskal's Algorithm

Goal: Find a Minimum Spanning Tree (MST) in a connected, weighted undirected graph.

Idea:

```
1. Start with an empty forest A (each vertex is its own tree).
2. Sort all edges in non-decreasing order of weight.
3. For each edge (u, v), if u and v are in different trees (i.e., no cycle is formed), add the edge to A.
4. Use a disjoint-set (Union-Find) data structure to efficiently check and merge trees.
```

```
1: procedure KRUSKAL( $G, w$ )
2:    $A \leftarrow \emptyset$ 
3:   for each  $v \in G.V$  do
4:     MAKE-SET( $v$ )
5:   end for
6:   sort the edges of  $G.E$  into non-decreasing order by weight w
7:   for each  $(u, v) \in$  sorted edge list do
8:     if FIND-SET( $u$ )  $\neq$  FIND-SET( $v$ ) then
9:        $A \leftarrow A \cup \{(u, v)\}$ 
10:      UNION( $u, v$ )
11:    end if
12:  end for
13:  return A
14: end procedure
```

Runtime:  $\Theta(|E| \log |E|)$  due to sorting, plus nearly linear time for Union-Find operations (with union by rank and path compression).

Edge Disjoint Paths using Max Flow

You can use the Max-Flow algorithm to find all edge-disjoint paths from a source to a sink by assigning a capacity of 1 to every edge and running Ford-Fulkerson. The maximum flow value will be equal to the number of edge-disjoint  $s$ - $t$  paths.

Bellman-Ford Algorithm

Goal: Compute shortest paths from a single source  $s$  to all other vertices in a weighted graph  $G = (V, E)$ , allowing negative edge weights.

Key Idea:

- Relax all edges repeatedly (up to  $|V| - 1$  times).
- After that, check for negative-weight cycles: if we can still relax an edge, a negative cycle exists.

Initialization:

```
1: procedure INIT-SINGLE-SOURCE( $G, s$ )
2:   for each  $v \in G.V$  do
3:      $v.d \leftarrow \infty$ 
4:      $v.\pi \leftarrow$  NIL
5:   end for
6:    $s.d \leftarrow 0$ 
7: end procedure
```

Relaxation:

```
1: procedure RELAX( $u, v, w$ )
2:   if  $v.d > u.d + w(u, v)$  then
3:      $v.d \leftarrow u.d + w(u, v)$ 
4:      $v.\pi \leftarrow u$ 
5:   end if
6: end procedure
```

Main Algorithm:

```
1: procedure BELLMAN-FORD( $G, w, s$ )
2:   INIT-SINGLE-SOURCE( $G, s$ )
3:   for i ← 1 to |G.V| - 1 do
4:     for each edge  $(u, v) \in G.E$  do
5:       RELAX( $u, v, w$ )
6:     end for
7:   end for
8:   for each edge  $(u, v) \in G.E$  do
9:     if  $v.d > u.d + w(u, v)$  then
10:      return false
11:    end if
12:  end for
13:  return true
14: end procedure
```

Runtime:  $\Theta(|V||E|)$

Handles: Negative weights (but no negative cycles)

Dijkstra's Algorithm

Goal: Compute the shortest paths from a single source  $s$  to all other vertices in a weighted graph  $G = (V, E)$  with non-negative edge weights.

Key Idea:

- Greedyly grow a set  $S$  of vertices with known shortest paths.
- At each step, pick the vertex  $u \notin S$  with the smallest tentative distance  $u, d$ .
- Relax all edges  $(u, v)$  from  $u$  to update distance estimates.

Pseudocode:

```
1: procedure DIJKSTRA( $G, w, s$ )
2:   INIT-SINGLE-SOURCE( $G, s$ )
3:    $S \leftarrow \emptyset$ 
4:    $Q \leftarrow G.V$  ▷ insert all vertices into priority queue Q
5:   while  $Q \neq \emptyset$  do
6:      $u \leftarrow$  EXTRACT-MIN( $Q$ )
7:      $S \leftarrow S \cup \{u\}$ 
8:     for each  $v \in \text{Adj}[u]$  do
9:       RELAX( $u, v, w$ )
10:    end for
11:  end while
12: end procedure
```

Runtime:  $\Theta(|E| \log |V|)$

Hash Tables

Hash tables are a data structure that use a function  $h(k)$  mapping keys to indices in the range 1 to  $p$ , such that each element is stored at index  $h(k)$ . Collisions are managed using chaining (linked lists), leading to:

- Insertion  $\mathcal{O}(1)$
- Search  $\mathcal{O}(N + E)$  (in worst-case)
- Deletion  $\mathcal{O}(1)$
- Collisions expected for  $m$  entries and  $n$  insertions (uniformly random hash function):  $\frac{n^2}{m}$

The Hiring Problem

Problem: We interview  $n$  candidates for a position, one by one in random order. After each interview, we decide immediately whether to hire the candidate. We want to compute the expected number of times we hire someone (i.e., when they are better than all previous candidates).

Indicator Random Variable: Given a sample space and an event  $A$ , the indicator random variable for  $A$  is defined as:

$$I\{A\} = \begin{cases} 1 & \text{if } A \text{ occurs,} \\ 0 & \text{if } A \text{ does not occur.} \end{cases}$$

Then:

$$\mathbb{E}[I\{A\}] = P(A)$$

Expected Number of Hires: Let  $X$  be the total number of hires. Define  $X = \sum_{i=1}^n I_i$ , where  $I_i = 1$  if the  $i$ -th candidate is hired (i.e., better than all previous  $i - 1$ ), and 0 otherwise.

$$\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[I_i] = \sum_{i=1}^n \frac{1}{i} = H_n$$

Conclusion: The expected number of hires is  $\Theta(\log n)$ , even though there are  $n$  candidates.

Quick Sort

Quick Sort is a divide-and-conquer algorithm with the following steps:

```
1. Divide: Partition  $A[p, \dots, r]$  into two (possibly empty) subarrays  $A[p, \dots, q-1]$  and  $A[q, \dots, r]$  such that each element in the first subarray is  $\leq A[q]$  and each element in the second subarray is  $\geq A[q]$ .
2. Conquer: Recursively sort the two subarrays by calling Quick Sort on them.
3. Combine: No work is needed to combine the subarrays since the sorting is done in-place.
```

Partition( $A, p, r$ ):

```
1: procedure PARTITION( $A, p, r$ )
2:    $x \leftarrow A[r]$ 
3:    $i \leftarrow p - 1$ 
4:   for  $j \leftarrow p$  to  $r - 1$  do
5:     if  $A[j] \leq x$  then
6:        $i \leftarrow i + 1$ 
7:       exchange  $A[i]$  with  $A[j]$ 
8:     end if
9:   end for
10:  exchange  $A[i + 1]$  with  $A[r]$ 
11:  return  $i + 1$ 
12: end procedure
```

Quick-Sort( $A, p, r$ ):

```
1: procedure QUICK-SORT( $A, p, r$ )
2:   if  $p < r$  then
3:      $q \leftarrow$  PARTITION( $A, p, r$ )
4:     QUICK-SORT( $A, p, q - 1$ )
5:     QUICK-SORT( $A, q + 1, r$ )
6:   end if
7: end procedure
```

Randomized-Partition( $A, p, r$ ):

```
1: procedure RANDOMIZED-PARTITION( $A, p, r$ )
2:    $i \leftarrow$  RANDOM( $p, r$ )
3:   exchange  $A[r]$  with  $A[i]$ 
4:   return PARTITION( $A, p, r$ )
5: end procedure
```

Randomized-Quick-Sort( $A, p, r$ ):

```
1: procedure RANDOMIZED-QUICK-SORT( $A, p, r$ )
2:   if  $p < r$  then
3:      $q \leftarrow$  RANDOMIZED-PARTITION( $A, p, r$ )
4:     QUICK-SORT( $A, p, q - 1$ )
5:     QUICK-SORT( $A, q + 1, r$ )
6:   end if
7: end procedure
```

Random Runtime:  $\Theta(N \log N)$

Worst Runtime:  $\Theta(N^2)$

Randomized caching

Each page is marked (if used recently) or unmarked.

- On miss, evict random unmarked page.
- Competitive ratio:  $2H(k) \approx \mathcal{O}(\log k)$  (nearly optimal, no randomized algorithm can beat  $H(k)$ ).

Online Algorithms

An online algorithm processes its input piece-by-piece in a serial fashion, i.e., it does not have access to the entire input from the start. Instead, it must make decisions based only on the current and past inputs without knowledge of future inputs.

Characteristics:

- Decisions are made in real-time.
- Cannot revise past decisions once new input arrives.
- Often evaluated using competitive analysis, comparing performance to an optimal offline algorithm.

Competitive Ratio: If  $\text{Online}$  is the cost incurred by the online algorithm and  $\text{Cost}_{\text{opt}}$  is the cost incurred by an optimal offline algorithm, then the competitive ratio is defined as:

$$\text{Competitive Ratio} = \max \frac{\text{Cost}_{\text{online}}}{\text{Cost}_{\text{opt}}}$$

An algorithm is said to be  $r$ -competitive if this ratio is at most  $r$  for all inputs.

Weighted Majority Algorithm

The Weighted Majority Algorithm (WMA) is an online learning algorithm that maintains a set of "experts" (prediction strategies), each assigned a weight. The algorithm predicts based on a weighted vote of the experts and penalizes those who make incorrect predictions.

Setup:

- $n$  experts, each with an initial weight  $w_i \leftarrow 1$ .
- At each time step  $t$ , each expert  $i$  makes a prediction.
- The algorithm makes its own prediction based on a weighted majority.
- After the outcome is revealed, experts that predicted incorrectly are penalized by multiplying their weight by a factor  $\beta \in (0, 1)$ .

Guarantees: If there is an expert that makes at most  $m$  mistakes, then the number of mistakes made by the algorithm is at most:

$$M \leq (1 + \log n) \cdot m$$

(up to constant factors depending on  $\beta$ )

Use cases: Binary prediction problems, stock forecasting, game playing.

Hedge Algorithm

The Hedge Algorithm generalizes Weighted Majority to handle real-valued losses and probabilistic predictions. It is used in adversarial multi-armed bandit and online optimization settings.

Setup:

- $n$  actions (or experts), each with weight  $w_i^{(t)}$  at round  $t$ .
- At each time step, the algorithm picks a probability distribution  $p^{(t)}$  over actions, where:

$$p_i^{(t)} = \frac{w_i^{(t)}}{\sum_j w_j^{(t)}}$$

After observing losses  $\ell_i^{(t)} \in [0, 1]$ , weights are updated as:

$$w_i^{(t+1)} = w_i^{(t)} \cdot e^{-\eta \ell_i^{(t)}}$$

where  $\eta > 0$  is the learning rate.

Guarantees: For any expert  $i$ , the regret after  $T$  rounds is bounded by:

$$\text{Regret} \leq \eta T + \frac{\log n}{\eta}$$

Setting  $\eta = \sqrt{\frac{\log n}{T}}$  gives regret of order  $\mathcal{O}(\sqrt{T \log n})$ .

Use cases: Adversarial learning, portfolio selection, online convex optimization.