Algorithm runtime analysis and computational tractability

As soon as an Analytic Engine exists, it will necessarily guide the future course of the science. Whenever any result is sought by its aid, the question will arise - By what course of calculation can these results be arrived at by the machine in the shortest time? - *Charles Babbage*

Charles Babbage (1864)  
Analytic Engine (schematic)
Time Complexity of an Algorithm

How do we measure the **complexity** (time, space requirements) of an algorithm.

The **size** of the problem: an integer \( n \)
- # inputs (for sorting problem)
- # digits of input (for the primality problem)
- sometimes more than one integer
  - Knapsack problem:
    - \( C \): capacity of the knapsack,
    - \( N \): number of objects

We want to characterize the running time of an algorithm for increasing problem sizes by a function \( T(n) \)
Units of time

1 microsecond?

1 machine instruction?

# of code fragments that take constant time?
Units of time

1 microsecond?

   no, too specific and machine dependent

1 machine instruction?

   no, still too specific and machine dependent

# of code fragments that take constant time?

   yes

# what kind of instructions take constant time?

   arithmetic op, memory access, finite combination of these
unit of space

bit?

int?
bit?

very detailed but sometimes necessary

int?

nicer, but dangerous: we can code a whole program or array (or disk) in one arbitrary sized int, so we have to be careful with space analysis (take value ranges into account when needed). Better to think in terms of machine words

i.e. fixed size, e.g. 64, collections of bits
Worst-Case Analysis

Worst case running time.

A bound on largest possible running time of algorithm on inputs of size n.
- Generally captures efficiency in practice, but can be an overestimate.

Same for worst case space complexity
Average case running time. A bound on the average running time of algorithm on random inputs as a function of input size $n$. In other words: the expected number of steps an algorithm takes.

\[
\sum_{i \in I_n} P_i C_i
\]

- $P_i$: probability input $i$ occurs
- $C_i$: complexity given input $i$
- $I_n$: all possible inputs of size $n$

- Hard to model real instances by random distributions.
- Algorithm tuned for a certain distribution may perform poorly on other inputs.
- Often hard to compute.
**Why It Matters**

Table 2.1 The running times (rounded up) of different algorithms on inputs of increasing size, for a processor performing a million high-level instructions per second. In cases where the running time exceeds $10^{25}$ years, we simply record the algorithm as taking a very long time.

<table>
<thead>
<tr>
<th></th>
<th>$n$</th>
<th>$n \log_2 n$</th>
<th>$n^2$</th>
<th>$n^3$</th>
<th>$1.5^n$</th>
<th>$2^n$</th>
<th>$n!$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 10$</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>4 sec</td>
</tr>
<tr>
<td>$n = 30$</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>18 min</td>
<td>$10^{25}$ years</td>
</tr>
<tr>
<td>$n = 50$</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>11 min</td>
<td>36 years</td>
<td>very long</td>
</tr>
<tr>
<td>$n = 100$</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>1 sec</td>
<td>12,892 years</td>
<td>$10^{17}$ years</td>
<td>very long</td>
</tr>
<tr>
<td>$n = 1,000$</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>1 sec</td>
<td>18 min</td>
<td>very long</td>
<td>very long</td>
<td>very long</td>
</tr>
<tr>
<td>$n = 10,000$</td>
<td>&lt; 1 sec</td>
<td>&lt; 1 sec</td>
<td>2 min</td>
<td>12 days</td>
<td>very long</td>
<td>very long</td>
<td>very long</td>
</tr>
<tr>
<td>$n = 100,000$</td>
<td>&lt; 1 sec</td>
<td>2 sec</td>
<td>3 hours</td>
<td>32 years</td>
<td>very long</td>
<td>very long</td>
<td>very long</td>
</tr>
<tr>
<td>$n = 1,000,000$</td>
<td>1 sec</td>
<td>20 sec</td>
<td>12 days</td>
<td>31,710 years</td>
<td>very long</td>
<td>very long</td>
<td>very long</td>
</tr>
</tbody>
</table>
A definition of tractability: Polynomial-Time

**Brute force.** For many problems, there is a natural brute force search algorithm that checks every possible solution.
- Typically takes $2^n$ time or worse for inputs of size $n$.
- Unacceptable in practice.
  - Permutations, TSP

An algorithm is said to be polynomial if there exist constants $c > 0$ and $d > 0$ such that on every input of size $n$, its running time is bounded by $c n^d$ steps.

- What about an $n \log n$ algorithm?
Worst-Case Polynomial-Time

On the one hand:

- Possible objection: Although $6.02 \times 10^{23} \times n^{20}$ is technically poly-time, it would be useless in practice, (e.g. $n=10$)
- In practice, the poly-time algorithms that people develop typically have low constants and low exponents.
- Breaking through the exponential barrier of brute force typically exposes some crucial structure of the problem.

On the other:

- Some exponential-time (or worse) algorithms are widely used because the worst-case (exponential) instances seem to be rare.
  - simplex method solving linear programming problems
Characterizing algorithm running times

Suppose that algorithm A has a running time bounded by

\[ T(n) = 1.62 \, n^2 + 3.5 \, n + 8 \]

- There is more detail than is useful
  - We want to quantify running time in a way that will allow us to identify broad classes of algorithms

- I.e., we only care about Orders of Magnitude
  - In this case: \( T(n) = O(n^2) \)
Asymptotic Growth Rates
Upper bounds

Recap from cs220:

T(n) is $O(f(n))$ if there exist constants $c > 0$ and $n_0 \geq 0$ such that

for all $n \geq n_0 : T(n) \leq c \cdot f(n)$

Example: $T(n) = 32n^2 + 16n + 32$.

- $T(n)$ is $O(n^2)$
- BUT ALSO: $T(n)$ is $O(n^3)$, $T(n)$ is $O(2^n)$.

There are many possible upper bounds for one function! We always look for the best (lowest) upper bound, but it is not always easy to establish.
Expressing Lower Bounds

Big O doesn’t always express what we want:

Any comparison-based sorting algorithm requires at least \( c(n \log n) \) comparisons, for some constant \( c \).

- Use \( \Omega \) for **lower bounds.**

\[ T(n) \text{ is } \Omega(f(n)) \text{ if there exist constants } c > 0 \text{ and } n_0 \geq 0 \text{ such that for all } n \geq n_0 \colon T(n) \geq c \cdot f(n) \]

**Example:** \( T(n) = 32n^2 + 16n + 32. \)

- \( T(n) \) is \( \Omega(n^2), \Omega(n) \).
Tight Bounds

T(n) is $\Theta(f(n))$ if $T(n)$ is both $O(f(n))$ and $\Omega(f(n))$.

Example: $T(n) = 32n^2 + 17n + 32$.
  - $T(n)$ is $\Theta(n^2)$.

If we show that the running time of an algorithm is $\Theta(f(n))$, we have closed the problem and found a bound for the problem and its algorithm solving it.
These measures were introduced in CS220.
Recap: rules of the game

- We are giving you extra time and trials for quiz 1 - Because we want you to be ready for the material in this class

- You need to know, and be comfortable with, the Pre-requisitate material

- You are responsible for all the material in the class
  - Readings from the text
  - Slides
  - Exercises
  - Tutorial
Excursion: heap sort and priority queues

**Priority Queue:** data structure that maintains a set $S$ of elements.

Each element $v$ in $S$ has a key $\text{key}(v)$ that denotes the priority of $v$.

Priority Queue provides support for adding, deleting elements, selection / extraction of smallest (Min prioQ) or largest (Max prioQ) key element, changing key value.
Applications

E.g. used in managing real time events where we want to get the earliest next event and events are added / deleted on the fly.

Sorting
- build a prioQ
- Iteratively extract the smallest element

PrioQs can be implemented using heaps
Heaps

Heap: array representation of a complete binary tree

- every level is completely filled except the bottom level: filled from left to right
- Can compute the index of parent and children: WHY?
  - parent(i) = floor((i-1)/2)
  - leftChild(i) = 2i+1
  - rightChild(i) = 2(i+1)

Max Heap property:
  for all nodes i>0: A[parent(i)] >= A[i]
  Max heaps have the max at the root

Min heaps have the min at the root
Heapify($A, i, n$)

To create a heap at index $i$, assuming left($i$) and right($i$) are heaps, bubble $A[i]$ down: swap with max child until heap property holds.

heapify($A, i, n$):
  # precondition
  # n is the size of the heap
  # tree left($i$) and tree right($i$) are heaps

  .......

  # postcondition: tree $A[i]$ is a heap
Swapping down enforces (max) heap property at the swap location:

\[
\begin{array}{c}
\text{new} \\
/ \ \\
\text{y} \\
\text{x}
\end{array} \rightarrow 
\begin{array}{c}
\text{x} \\
/ \\
\text{y} \\
\text{new}
\end{array}
\]

\[
\text{new} < x \quad \text{and} \quad y < x: \quad x > y \quad \text{and} \quad x > \text{new}
\]

\[
\text{swap}(x, \text{new})
\]

Are we done now?

NO! When we have swapped we need to carry on checking whether \( \text{new} \) is in heap position. We stop when that is the case.
Heap Extract

Heap extract:
Delete (and return) root

**Step 1:** replace root with last array element to keep completeness

**Step 2:** reinstate the heap property
Which element does not necessarily have the heap property?

How can it be fixed? Complexity?

heapify the root $O\left(\log n\right)$

Swap *down:* swap with maximum (maxheap), minimum (minheap) child as necessary, until in place.

Sometimes called bubble down

Correctness based on the fact that we started with a heap, so the children of the root are heaps
Heap Insert

Step 1: put a new value into first open position (maintaining completeness), i.e. at the end of the array, but now we potentially violated the heap property, so:

Step 2: bubble up

- Re-enforcing the heap property
- Swap with parent, if new value > parent, until in the right place.

- The heap property holds for the tree below the new value, when swapping up. WHY? We only compared the new element to the parent, not to the sibling!
Swapping up enforces heap property for the sub tree below the new, inserted value:

If \( new > x \) swap\((x, new)\)  
\[
\begin{align*}
\text{if } (\text{new} > x) & \text{ swap}(x, \text{new}) \\
\text{x} & \text{ new} \\
\text{y} & \text{ new} \\
\text{y} & x
\end{align*}
\]

\( x > y \), therefore \( \text{new} > y \)
Building a heap

heapify performs at most $\lg n$ swaps

why?  what is $n$?

buildheap:  builds a heap out of an array:

- the leaves are all heaps  WHY?
- heapify backwards starting at last internal node

WHY backwards?
WHY last internal node?
which node is that?
LERT'S DO THE BUILDHEAP!

[4, 8, 7, 2, 14, 1]
Complexity buildheap

Suggestions? ...
Initial thought: $O(n \log n)$, but

half of the heaps are height 0
quarter are height 1
only one is height $\log n$

It turns out that $O(n \log n)$ is not tight!
complexity buildheap

height  max #swaps ?
0
1
2
3
max #swaps, see a pattern? (What kind of growth function do you expect?)

0

1

2*1+2 = 4

3

2*4+3 = 11
complexity buildheap

<table>
<thead>
<tr>
<th>height</th>
<th>max #swaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 = $2^1 - 2$</td>
</tr>
<tr>
<td>1</td>
<td>1 = $2^2 - 3$</td>
</tr>
<tr>
<td>2</td>
<td>$2 \times 1 + 2 = 4 = 2^3 - 4$</td>
</tr>
<tr>
<td>3</td>
<td>$2 \times 4 + 3 = 11 = 2^4 - 5$</td>
</tr>
</tbody>
</table>
Conjecture:
height = h
max #swaps = 2^{h+1}-(h+2)

Proof: induction
base?
step:
height = (h+1)
max #swaps:
\[ 2^h(2^{h+1}-(h+2))+(h+1) \]
\[ = 2^{h+2}-2h-4+h+1 \]
\[ = 2^{h+2}-(h+3) \]
\[ = 2^{(h+1)+1}-(((h+1)+2)) \]

n nodes $\rightarrow \Theta(n)$ swaps
See it the Master theorem way

\[ T(n) = 2 \cdot T(n/2) + \lg n \]

Master theorem \[ \Theta(n^{\lg_2 2}) = \Theta(n) \]
Heapsort, complexity

heapsort(A):
  buildheap(A)  # O( n )
  for i = n-1 downto 1:
    # put max at end array

    # max is removed from heap
    n=n-1

    # reinstate heap property  # *( lg n )

- heapify: Θ(lgn)
- heapExtract: Θ(lg n)
- buildheap: Θ(n)
- heapsort: Θ(n lg n)
- space: in place: Θ(n)
DO THE HEAPSORT, DO IT, DO IT!

1. 14
2. 8
3. 7
4. 2
5. 4
6. 1

1. 8
2. 4
3. 1
4. 7
5. 2
6. 1

1. 7
2. 8
3. 14
4. 2
5. 4
6. 1

1. 4
2. 7
3. 8
4. 14
5. 2
6. 4
7. 7
8. 8
9. 14

1. 2
2. 4
3. 7
4. 8
5. 14
6. 1
7. 2
8. 4
9. 7
10. 8
11. 14
How **not** to heapExtract, heapInsert

# These "snail" implementations are NOT preserving the algorithm
# complexity of extractMin: log n and insert: log n and are therefore
# **INCORRECT!** from a complexity point of view (even though they are
# functionally correct). Remember one of the goals of our course:
# implementing the algorithms maintaining the analyzed complexity
# What are their complexities?

```python
def snailExtractMin(A):
    n = len(A)
    if n == 0:
        return None
    min = A[0]
    A[0]=A[n-1]
    A.pop()
    buildHeap(A)  #  O(n)
    return min

def snailInsert(A,v):
    A.append(v)
    buildHeap(A)  #   O(n)
```

Priority Queues

Heaps can be used to implement priority queues:

- each value associated with a key
- max priority queue S has operations that maintain the heap property of S
  - \text{max}(S) returning max element
  - \text{Extract-max}(S) extracting and returning max element
  - \text{increase key}(S,x,k) increasing the key value of x
  - \text{insert}(S,x)
    - put x at end of S
    - bubble x up in place
Transitivity.
- If \( f = O(g) \) and \( g = O(h) \) then \( f = O(h) \).
- If \( f = \Omega(g) \) and \( g = \Omega(h) \) then \( f = \Omega(h) \).
- If \( f = \Theta(g) \) and \( g = \Theta(h) \) then \( f = \Theta(h) \).

Additivity.
- If \( f = O(h) \) and \( g = O(h) \) then \( f + g = O(h) \).
- If \( f = \Omega(h) \) and \( g = \Omega(h) \) then \( f + g = \Omega(h) \).
- If \( f = \Theta(h) \) and \( g = \Theta(h) \) then \( f + g = \Theta(h) \).
Polynomials. $a_0 + a_1 n + ... + a_d n^d$ is $O(n^d)$ if $a_d > 0$.

Polynomial time. Running time is $O(n^d)$ for some constant $d$.

Logarithms. $\log_a n$ is $O(\log_b n)$ for any constants $a, b > 0$.

\[
\text{can avoid specifying the base}
\]

\[
\text{for every } x > 0, \log n \text{ is } O(n^x).
\]

\[
\log \text{ grows slower than any polynomial}
\]

Combinations. Merge sort, Heap sort $O(n \log n)$

Exponentials. For every $r > 1$ and every $d > 0$, $n^d$ is $O(r^n)$.

\[
\text{exponential grows faster than any polynomial}
\]
Problems have lower bounds

A problem has a lower bound \( \Omega(f(n)) \), which means:

Any algorithm solving this problem takes at least \( \Omega(f(n)) \) steps.

We can often show that an algorithm has to "touch" all elements of a data structure, or produce a certain sized output. This then gives rise to an easy lower bound.

Sometimes we can prove better (higher, stronger) lower bounds (eg Searching and Sorting (cs420)).
Closed / open problems

Problems have lower bounds, algorithms have upper bounds. A closed problem has a lower bound $\Omega(f(n))$ and at least one algorithm with upper bound $O(f(n))$

- Example: sorting is $\Omega(n \log n)$ and there are $O(n \log n)$ sorting algorithms.
  
  To show this, we need to reason about lower bounds of problems (CS420)

An open problem has lower bound < upper bound

- Example: matrix multiplication (multiply two $n \times n$ matrices).
  
  - Takes $\Omega(n^2)$ why?
  
  - Naïve algorithm: $O(n^3)$
  
  - Coppersmith-Winograd algorithm: $O(n^{2.376})$
A Survey of Common Running Times
A single line of code that involves “simple” expressions, e.g.:
- Arithmetical operations (+,-,*,/) for fixed size inputs
- assignments (x = simple expression)
- conditionals with simple sub-expressions
- function calls (excluding the time spent in the called function)
Logarithmic time

Example of a problem with $O(\log(n))$ bound:

- binary search

How did we get that bound?
Guessing game

I have a number between 0 and 63
How many (Y/N) questions do you need to find it?

What's the number?

What (kind of) questions would you ask?
Guessing game

I have a number between 0 and 63
How many (Y/N) questions do you need to find it?

is it >= 32    N
is it >= 16    Y
is it >= 24    N
is it >= 20    N
is it >= 18    Y
is it >= 19    Y

What’s the number?
  19

Take N=0 and Y=1, what is 010011 ?
log(n) and algorithms

When in each step of an algorithm we halve the size of the problem then it takes $\log_2 n$ steps to get to the base case.

We often use $\log(n)$ when we should use $\lfloor \log(n) \rfloor$. That's OK since $\lfloor \log(n) \rfloor$ is $\Theta(\log(n))$.

Similarly, if we divide a problem into $k$ equal parts the number of steps is $\log_k n$. For the purposes of big-O analysis it doesn't matter since $\log_a n$ is $O(\log_b n)$.
Logarithms

definition:
\[ b^x = a \implies x = \log_b a, \text{ eg } 2^3 = 8, \log_2 8 = 3 \]
\[ b^{\log_b a} = a \quad \log_b b = 1 \quad \log 1 = 0 \]

- \[ \log(x \cdot y) = \log x + \log y \text{ because } b^x \cdot b^y = b^{x+y} \]
- \[ \log(x/y) = \log x - \log y \]
- \[ \log x^a = a \log x \]
- \[ \log x \text{ is a 1-to-1 monotonically (slow) growing function} \]
  \[ \log x = \log y \iff x = y \]

- \[ \log_a x = \log_b x / \log_b a \]
- \[ y^{\log x} = x^{\log y} \]
\[ \log_a x = \log_b x / \log_b a \]

\[ b^{\log_b x} = x = a^{\log_a x} = b^{(\log_b a)(\log_a x)} \]

\[ \log_b x = (\log_b a)(\log_a x) \]

\[ \log_a x = \log_b x / \log_b a \]

therefore \( \log_a x = O(\log_b x) \) for any \( a \) and \( b \)
\[ y^{\log x} = x^{\log y} \]

\[ x^{\log_b y} = \]

\[ y^{\log_y x \log_b y} = \]

\[ y^{(\log_b x / \log_b y) \log_b y} = \]

\[ y^{\log_b x} = \]
Combinations of functions /code fragments

AdditiveTheorem:

Suppose that $f_1(x)$ is $O(g_1(x))$ and $f_2(x)$ is $O(g_2(x))$. Then $(f_1 + f_2)(x)$ is $O(\max(g_1(x), g_2(x)))$.

Sequences of code are additive in complexity:

```c
int c = 0;
for(int i=0; i<n; i++)
    c++;
for(int j=0; j<m; j++)
    c++;
```

Complexity?

What is counting the complexity?
Combinations of functions /code fragments

Multiplicative Theorem:
Suppose that $f_1(x)$ is $O(g_1(x))$ and $f_2(x)$ is $O(g_2(x))$.
Then $(f_1f_2)(x)$ is $O(g_1(x)g_2(x))$.

Nested code is multiplicative in complexity
for(int i=0; i<n; i++)
  for(int j=0; j<m; j++)
    c++;

Complexity?

BUT, be careful with nests where the inner loop depends outer loop:
int b = n;
while(b>0){
  b/=2;
  for(int i=0; i<b; i++)
    c++;
}
Recursive Code

Draw the call tree, and assert the number of nodes in the tree and their individual complexity, as a function of $n$. 
Recursive Code

Draw the call tree, and assert the number of nodes in the tree and their individual complexity, as a function of $n$.

```java
public int divCo(int n){
    if(n<=1)
        return 1;
    else
        return 1 + divCo(n-1) + divCo(n-1);
}
```

How many recursive calls?

How much work per call?
What is the role of “return 1” and return 1+...” ?
So what does this function count?

Big O complexity?
Linear time. Running time is proportional to the size of the input.

Computing the maximum. Compute maximum of n numbers $a_1, \ldots, a_n$.

\begin{verbatim}
max ← $a_1$
for $i = 2$ to $n$ {
    if ($a_i > max$)
        $max ← a_i$
}
\end{verbatim}

Also $\Theta(n)$?
**Linear Time: O(n)**

**Merge.** Combine two sorted lists $A = a_1, a_2, \ldots, a_n$ with $B = b_1, b_2, \ldots, b_n$ into a single sorted list.

\[
\text{i = 1, j = 1} \\
\text{while (both lists are nonempty) \{} \\
\quad \text{if (} a_i \leq b_j \text{) append } a_i \text{ to output list and increment } i \\
\quad \text{else append } b_j \text{ to output list and increment } j \\
\text{\}} \\
\text{append remainder of nonempty list to output list}
\]

**Claim.** Merging two lists of size n takes $O(n)$ time.
Linear Time: $O(n)$

Polynomial evaluation. Given

$$A(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_0 \quad (a_n \neq 0)$$

Evaluate $A(x)$

How not to do it:

$$a_n \exp(x, n) + a_{n-1} \exp(x, n-1) + ... + a_1 x + a_0$$

Why not?
How to do it: Horner's rule

\[ a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x^1 + a_0 = \]

\[ (a_n x^{n-1} + a_{n-1} x^{n-2} + \ldots + a_1) x + a_0 = \ldots = \]

\[ (\ldots (a_n x + a_{n-1}) x + a_{n-2}) x \ldots + a_1) x + a_0 \]

\[ y = a[n] \]

\[ \text{for } (i=n-1; i>=0; i--) \]

\[ y = y \times x + a[i] \]
Polynomial evaluation using Horner: complexity

Lower bound: \( \Omega(n) \) because we need to access each \( a[i] \) at least once

Upper bound: \( O(n) \)

Closed problem!

But what if \( A(x) = x^n \)
A(x) = x^n

Recurrence:
\[ x^{2n} = x^n \times x^n \]
\[ x^{2n+1} = x \times x^{2n} \]

```python
def pwr(x, n):
    if (n == 0):
        return 1
    if odd(n):
        return x * pwr(x, n-1)
    else:
        a = pwr(x, n/2)
        return a * a
```

Complexity?
A glass-dropping experiment

- You are testing a model of glass jars, and want to know from what height you can drop a jar without it breaking. You can drop the jar from heights of 1,…,n foot heights. Higher means faster means more likely to break.
- You want to minimize the amount of work (number of heights you drop a jar from). Your strategy would depend on the number of jars you have available.
  - If you have a single jar:
    - do linear search (O(n) work).
  - If you have an unlimited number of jars:
    - do binary search (O(log n) work)
  - Can you design a strategy for the case you have 2 jars, resulting in a bound that is strictly less than O(n)?
Often arises in divide-and-conquer algorithms like mergesort.

```python
mergesort(A) :
    if len(A) <= 1 return A
    else return merge(mergesort(left half(A)),
                        mergesort(right half(A)))
```
Merge Sort - Divide

\{7,3,2,9,1,6,4,5\}

\{7,3,2,9\} \quad \{1,6,4,5\}

\{7,3\} \quad \{2,9\}

\{7\} \quad \{3\} \quad \{2\} \quad \{9\}

\{1,6\} \quad \{4,5\}

\{1\} \quad \{6\} \quad \{4\} \quad \{5\}
mergesort(A):
    if len(A) <= 1 return A
    else return merge(mergesort(left half(A)),
                      mergesort(right half(A)))
Quadratic Time: $O(n^2)$

Quadratic time example. Enumerate all pairs of elements.

Closest pair of points. Given a list of $n$ points in the plane $(x_1, y_1), \ldots, (x_n, y_n)$, find the pair that is closest.

$O(n^2)$ solution. Try all pairs of points.

\[
\begin{align*}
\min & \leftarrow (x_1 - x_2)^2 + (y_1 - y_2)^2 \\
& \text{for } i = 1 \text{ to } n \{ \\
& \quad \text{for } j = i+1 \text{ to } n \{ \\
& \quad \quad d \leftarrow (x_i - x_j)^2 + (y_i - y_j)^2 \\
& \quad \quad \text{if } (d < \min) \\
& \quad \quad \quad \min \leftarrow d \\
& \quad \} \\
& \} \\
\end{align*}
\]

Remark. $\Omega(n^2)$ seems inevitable, but . . . .
Cubic Time: $O(n^3)$

Example 1: Matrix multiplication

Tight?

Example 2: Set disjoint-ness. Given $n$ sets $S_1, \ldots, S_n$ each of which is a subset of $1, 2, \ldots, n$, is there some pair of these which are disjoint?

$O(n^3)$ solution. For each pairs of sets, determine if they are disjoint.

```plaintext
foreach set $S_i$ {
    foreach other set $S_j$ {
        foreach element $p$ of $S_i$ {
            determine whether $p$ also belongs to $S_j$
        }
        if (no element of $S_i$ belongs to $S_j$)
            report that $S_i$ and $S_j$ are disjoint
    }
}
```

what do we need for this to be $O(n^3)$?
Largest interval sum

Given an array $A[0], \ldots, A[n - 1]$, find indices $i, j$ such that the sum $A[i] + \ldots + A[j]$ is maximized.

Naïve algorithm:

$$\text{maximum_sum} = -\infty$$

for $i$ in range($n - 1$):
  for $j$ in range($i, n$):
    $\text{current_sum} = A[i] + \ldots + A[j]$
    if $\text{current_sum} \geq \text{maximum_sum}$:
      $\text{maximum_sum} = \text{current_sum}$
      save the values of $i$ and $j$

big $O$ bound?

Can we do better?

Example:

$A = [2, -3, 4, 2, 5, 7, -10, 8, 12]$
Polynomial Time: $O(n^k)$ Time

**Independent set of size $k$.** Given a graph, are there $k$ nodes such that no two are joined by an edge?

$k$ is a constant

**$O(n^k)$ solution.** Enumerate all subsets of $k$ nodes.

```plaintext
foreach subset $S$ of $k$ nodes {
    check whether $S$ is an independent set
    if (S is an independent set)
        report $S$ is an independent set
}
```

- Check whether $S$ is an independent set = $O(k^2)$.
- Number of $k$ element subsets = $\binom{n}{k} = \frac{n (n-1) (n-2) \cdots (n-k+1)}{k (k-1) (k-2) \cdots (2) (1)} \leq \frac{n^k}{k!}$
- $O(k^2 n^k / k!) = O(n^k)$.

poly-time for $k=17$, but not practical
Exponential Time

**Independent set.** Given a graph, what is the maximum size of an independent set?

**$O(n^2 2^n)$ solution.** Enumerate all subsets.

```plaintext
S* ← ∅
foreach subset S of nodes {
    check whether S in an independent set
    if (S is largest independent set seen so far)
        update S* ← S
}
```

For some problems (e.g. TSP) we need to consider all permutations. The factorial function (n!) grows much faster than $2^n$. 
Questions

1. Is $2^n \in O(3^n)$ ?
2. Is $3^n \in O(2^n)$
3. Is $2^n \in O(n!)$ ?
4. Is $n! \in O(2^n)$
Some problems (such as matrix multiply) have a polynomial complexity solution: an $O(n^p)$ time algorithm solving them. (p constant)

Some problems (such as Hanoi) take an exponential time to solve: $\Theta(p^n)$ (p constant)

For some problems we only have an exponential solution, but we don't know if there exists a polynomial solution. Trial and error algorithms are the only ones we have so far to find an exact solution, and if we would always make the right guess, these algorithms would take polynomial time.

We call these problems NP (non deterministic polynomial) We will discuss NP later.
Some NP problems

**TSP**: Travelling Salesman

given cities $c_1, c_2, \ldots, c_n$ and distances between all of these, find a minimal tour connecting all cities.

**SAT**: Satisfiability

given a boolean expression $E$ with boolean variables $x_1, x_2, \ldots, x_n$ determine a truth assignment to all $x_i$ making $E$ true
Back tracking

Back tracking searches (walks) a state space, at each choice point it guesses a choice.

In a leaf (no further choices) if solution found OK, else go back to last choice point and pick another move.

NP is the class of problems for which we can check in polynomial time whether it is correct (certificates, later)
Coping with intractability

NP problems become intractable quickly

TSP for 100 cities?

How would you enumerate all possible tours? How many?

Coping with intractability:

- Approximation: Find a nearly optimal tour
- Randomization: use a probabilistic algorithm using "coin tosses" (eg prime witnesses)