Learning Objectives

By the end of this worksheet, you will:

- Analyse the average-case running time of an algorithm.

---

1. **Average-case analysis.** Consider the following algorithm that we studied a few weeks ago. The input is an array A of length n, containing a list of n integers.

```
def has_even(lst: List[int]) -> int:
    n = len(lst)
    for i in range(n):
        if lst[i] % 2 == 0:
            return i

    return -1
```

We proved that the worst-case running time of this algorithm is $\Theta(n)$. In this problem we will analyse its *average-case* running time.

For this analysis, we will consider the sets of *binary* lists lst of length n, for each $n \in \mathbb{Z}^+$. That is, lst is a list of n integers, where each integer is either 0 or 1.

(a) For each $n \in \mathbb{Z}^+$, let $\mathcal{I}_n$ be the set of all binary lists of length n. Find an expression (in terms of n) for $|\mathcal{I}_n|$, the size of $\mathcal{I}_n$.

size of In = 2^n

(b) For each $n \in \mathbb{Z}^+$ and each $i \in \{0, 1, \ldots, n-1\}$, let $S_{n,i}$ denote the *set* of all binary lists lst of length n where the first 0 occurs in position i. More precisely, every list lst in $S_{n,i}$ satisfies the following two properties:

(i) $lst[i] = 0$.

(ii) for all $j \in \mathbb{N}$, if $j < i$ then $lst[j] = 1$.

For each i, $0 \leq i \leq n$, find an expression for $|S_{n,i}|$.

spaces after i = n - i - 1

S = 2^ (n- i - 1)

(c) Also, for each $n \in \mathbb{Z}^+$, let $S_{n,n}$ denote the set of binary lists of length n that do not contain a 0 at all. Find an expression for $|S_{n,n}|$.

1

(d) Give a brief argument (informal proof) that for every $n \in \mathbb{Z}^+$, each binary list of length n is in exactly one set $S_{n,i}$ (for some $i \in \{0, 1, \ldots, n\}$). That is, you're arguing that $S_{n,0}, S_{n,1}, \ldots, S_{n,n}$ form a *partition* of $\mathcal{I}_n$.

In = lists with 0, all lists without 0
Lists with a 0 in it, the index of first 0 = i, i = {0, 1, …, n-1}
List without a 0, [0,0,0,..,0] length n = Snn

The sets don't overlap. Unique position of
list in S.

(e) Assume that we calculate the running time of `has_even` by counting just the costs of Lines 4 and 7. Find an exact expression for the average runtime of this algorithm for this input set $\mathcal{I}_n$, in terms of n.
You should get a summation; do not simplify the summation in this part.

sum_{i=0}^{n} (i + 1)(S_n,i) /|In|
(n+1)1+ sum_{i=0}^{n-1} (i + 1)(2^(n - i - 1)) /2^n

T(n) = max{t(i): |i| = n}

let i' = i + 1
n+1 + sum_{i'=1}^{n}i'* 2^(n - i')

2^n

(f) Show that the average running time expression that you calculated is in $\mathcal{O}(1)$. You may use the fact that for all $x \in \mathbb{R}$, if $|x| < 1$, then $\displaystyle\sum_{i=1}^{\infty} ix^i = \frac{x}{(1-x)^2}$.