

# Computational Complexity

Lecture 12: Average-case complexity and Impagliazzo's Five Worlds

Ronald de Haan

me@ronalddehaan.eu

University of Amsterdam

March 12, 2026

- Subexponential-time complexity
- Exponential-Time Hypothesis (ETH)

## What will we do today?

- Average-case complexity
- One-way functions
- Impagliazzo's Five Worlds

- A problem  $L \subseteq \{0, 1\}^*$  can be solved in *worst-case running time*  $T(n)$  if there exists an algorithm  $A$  that solves  $L$  and that halts within time  $T(|x|)$  for each  $x \in \{0, 1\}^*$ .
- In other words, the worst-case running time  $T(n)$  is the maximum of the running times for all inputs of size  $n$ .

### Definition (distributional problems)

A *distributional problem*  $\langle L, \mathcal{D} \rangle$  consists of a language  $L \subseteq \{0, 1\}^*$  and a sequence  $\mathcal{D} = \{\mathcal{D}_n\}_{n \in \mathbb{N}}$  of probability distributions, where each  $\mathcal{D}_n$  is a probability distribution over  $\{0, 1\}^n$ .

## Definition (distP)

$\langle L, \mathcal{D} \rangle$  is in the class distP (also called: avgP) if there exists a deterministic TM  $\mathbb{M}$  that decides  $L$  and a constant  $\epsilon > 0$  such that for all  $n \in \mathbb{N}$ :

$$\mathbb{E}_{x \in_R \mathcal{D}_n} [ \text{time}_{\mathbb{M}}(x)^\epsilon ] \text{ is } O(n).$$

- The  $\epsilon$  is there for technical reasons—to invert a polynomial to  $O(n)$ .

## Definition (P-computable distributions)

A sequence  $\mathcal{D} = \{\mathcal{D}_n\}_{n \in \mathbb{N}}$  of distributions is *P-computable* if there exists a polynomial-time TM that, given  $x \in \{0, 1\}^n$ , computes:

$$\mu_{\mathcal{D}_n}(x) = \sum_{\substack{y \in \{0,1\}^n \\ y \leq x}} \mathbb{P}_{\mathcal{D}_n}[y],$$

where  $y \leq x$  if the number represented by the binary string  $y$  is at most the number represented by the binary string  $x$ .

### Definition (P-samplable distributions)

A sequence  $\mathcal{D} = \{\mathcal{D}_n\}_{n \in \mathbb{N}}$  of distributions is *P-samplable* if there exists a polynomial-time probabilistic TM  $\mathbb{M}$  such that for each  $n \in \mathbb{N}$ , the random variables  $\mathbb{M}(1^n)$  and  $\mathcal{D}_n$  are equally distributed.

### Definition (distNP)

A problem  $\langle L, \mathcal{D} \rangle$  is in distNP if  $L \in \text{NP}$  and  $\mathcal{D}$  is P-computable.

### Definition (sampNP)

A problem  $\langle L, \mathcal{D} \rangle$  is in sampNP if  $L \in \text{NP}$  and  $\mathcal{D}$  is P-samplable.

- The questions “distNP  $\stackrel{?}{=} \text{distP}$ ” and “sampNP  $\stackrel{?}{=} \text{distP}$ ” are average-case analogues of the question “NP  $\stackrel{?}{=} \text{P}$ ”

## Definition (one-way functions)

A polynomial-time computable function  $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$  is a *one-way function* if for every polynomial-time probabilistic TM  $\mathbb{M}$  there is a negligible function  $\epsilon : \mathbb{N} \rightarrow [0, 1]$  such that for every  $n \in \mathbb{N}$ :

$$\mathbb{P}_{\substack{x \in_{\mathbf{R}} \{0,1\}^n \\ y=f(x)}} [ \mathbb{M}(y) = x' \text{ such that } f(x') = y ] < \epsilon(n)$$

where a function  $\epsilon : \mathbb{N} \rightarrow [0, 1]$  is *negligible* if  $\epsilon(n) = \frac{1}{n^{\omega(1)}}$ , that is, for every  $c$  and sufficiently large  $n$ ,  $\epsilon(n) < \frac{1}{n^c}$ .

- Conjecture: there exist one-way functions (implying  $P \neq NP$ )
- OWFs can be used to create private-key cryptography

## Definition

An *encryption scheme* is a pair  $(E, D)$  of algorithms, each taking a key  $k$  and a message  $x$ , such that  $D_k(E_k(x)) = x$ .

The scheme is *perfectly secret*, for messages of length  $m$  and keys of length  $n$ , if for every pair  $x, x' \in \{0, 1\}^m$  of messages, the distributions  $E_{U_n}(x)$  and  $E_{U_n}(x')$  are identical.

The scheme is *computationally secure* if for every probabilistic polynomial-time algorithm  $A$ , there is a negligible function  $\epsilon : \mathbb{N} \rightarrow [0, 1]$  such that

$$\mathbb{P}_{\substack{k \in_{\mathbb{R}} \{0,1\}^n \\ x \in_{\mathbb{R}} \{0,1\}^m}} [ A(E_k(x)) = (i, b) \text{ s.t. } x_i = b ] < 1/2 + \epsilon(n).$$

- Suppose that OWFs exist. Then for every  $c \in \mathbb{N}$  there exists a computationally secure encryption scheme  $(E, D)$  using  $n$ -length keys for  $n^c$ -length messages.

Five possible situations regarding the status of various complexity-theoretic assumptions:

- Algorithmica
- Heuristica
- Pessiland
- Minicrypt
- Cryptomania

**Russell Impagliazzo.** *A personal view of average-case complexity.* In: Proceedings of the 10th Annual IEEE Conference on Structure in Complexity Theory, pp. 134–147, 1995.

- $P = NP$  (or  $NP \subseteq BPP$ )
- ▶ Say, SAT is linear-time solvable
- ▶ This is a computational utopia
- ▶ There exist efficient algorithms for creative tasks, e.g., writing proofs
- ▶ Essentially no cryptography possible (private-key nor public-key)

- $P \neq NP$ , but  $\text{distNP}, \text{sampNP} \subseteq \text{distP}$
- ▶ Breakthroughs of  $P = NP$  work almost all the time
- ▶ So cryptography breaks too

- $\text{distNP}, \text{sampNP} \not\subseteq \text{distP}$  (so  $P \neq \text{NP}$ )
- one-way functions do not exist
- ▶ No computational breakthroughs, and most cryptography schemes do not work

- One-way functions exist (so  $P \neq NP$  and  $\text{distNP} \not\subseteq \text{distP}$ )
- ▶ No “ $P = NP$ ”-type breakthroughs
- ▶ Private-key cryptography works
- ▶ All “highly structured” problems in NP, such as integer factoring, are solvable in polynomial-time
- ▶ Public-key cryptography might not work

- Factoring large integers takes exponential time on average (or a corresponding result for a similar problem)
- ▶ No general-purpose efficient algorithms ( $P \neq NP$ )
- ▶ Private-key and public-key cryptography works

- Five worlds:
  - Algorithmica – efficient general-purpose algorithms
  - Heuristica
  - Pessiland – worst of all worlds
  - Minicrypt
  - Cryptomania – all kinds of cryptography possible
- (Technically, these cases are not exhaustive—there are some “weirdland” scenarios, e.g., the case where  $\text{SAT} \in \text{P}$ , but the fastest algorithm takes time  $\Theta(n^{100})$ .)

- Average-case complexity
- One-way functions
- Impagliazzo's Five Worlds