COMP 590: Computational Complexity **Spring 2026**

Problem Set 3

Instructions: You **must** typeset your solution in LaTeX using the provided template. Please submit your problem set via Canvas. Include your name and the names of any collaborators at the top of your submission.

Name:
Collaborators:

Acknowledgment: Most problems come from the Arora Barak textbook and are reproduced here for convenience and typo fixes.

Problem 1: Steve's Class. Define **polyL** to be $\cup_{c>0}$**SPACE**$(\log^c n)$. Steve's Class **SC** (named in honor of Steve Cook) is defined to be the set of languages that can be decided by deterministic machines that run in polynomial time and $\log^c n$ space for some $c > 0$.

It is an open problem whether PATH $\in$ **SC**. Why does Savitch's Theorem not resolve this question? Is **SC** the same as **polyL** $\cap$ **P**?

Problem 2: P/poly Practice. Describe a *decidable* language in **P/poly** that is not in **P**.

Problem 3: Small Circuits, Small Space. Show that uniform $\mathbf{NC^1} \subseteq \mathbf{L}$. Why does this imply that **PSPACE** $\neq$ uniform $\mathbf{NC^1}$?

Problem 4: Improved Set Equality. In class we saw a protocol for checking if Alice and Bob hold identical sets $\{a_1, ..., a_n\}$ and $\{b_1, ..., b_n\}$. It was presented using a univariate version of polynomial identity testing.

(a) Show, given the numbers (a, n, p) in binary representation, how to compute $a^n \pmod p$ in polynomial time.

Hint: Use the binary representation of n and repeated squaring. The solution will take time $O(\log n)$. It might help to start by considering only the special case where n is a power of 2.

(b) Describe the set equality protocol again, but this time use a multivariate polynomial over $\mathbb{F}_p$ rather than a univariate polynomial, so the proof of correctness will rely on the Schwartz-Zippel lemma.

(c) Write down the asymptotic communication and computation costs for the parties in both this version of the protocol and the version from class, as well as the failure probability of each protocol. What are the benefits of using the multivariate version of the protocol?

(d) Suppose Alice and Bob have the additional restriction in your new set equality protocol that they can only send each other a small number of bits every day (e.g., they have a very limited data plan), forcing them to use a relatively small choice of modulus p_{small}.

If they want to get failure probability the same or smaller than the protocol with the original, larger p, on how many consecutive days will they need to run the protocol with p_{small}? How does the overall communication compare to just running the protocol once with the larger p?

Problem 5: Random Sampling. In this problem, we will show that one can efficiently simulate choosing a random number from 1 to N using coin tosses.

In particular, we will show that for every N and $p > 0$, there is a probabilistic algorithm A running in $\mathsf{poly}(\log N \log(1/p))$ time with output in $\{1, ..., N, ?\}$ such that
1) conditioned on not outputting ?, A's output is uniformly distributed in $\{1, ..., N\}$, and
2) the probability that A outputs ? is at most p.

(a) Show the algorithm.

Hint: start by coming up with a $\mathsf{poly}(\log N)$ time algorithm that only satisfies criterion (1).

(b) Demonstrate that your solution meets criteria (1) and (2).

(c) The largest prime number that can fit in 128 bits is $p = 2^{128} - 159$. In cryptographic applications where we cannot accept a ? outcome, we sometimes just sample 128 random bits, interpret that as a number, and take that number mod $2^{128} - 159$ as an element of $\mathbb{F}_p$.

While this is not exactly a uniformly random value in $\mathbb{F}_p$, it is close enough that it works for most applications. When we say "close enough," we mean that the *statistical distance* between this value and a truly uniformly random element of $\mathbb{F}_p$ is small, e.g., less than 2^{-100}.

For two distributions X and Y whose range is the finite set Ω, the statistical distance between X and Y is defined as

$$\begin{aligned}\Delta(X, Y) &= \max_{S \subseteq \Omega} \{|\Pr[X \in S] - \Pr[Y \in S]|\} \\ &= \frac{1}{2} \Sigma_{x \in \Omega} |\Pr[X = x] - \Pr[Y = x]|.\end{aligned}$$

What is the statistical distance between the uniform distribution over $\mathbb{F}_p$ and the sampling process described here? Why would this same trick not work if we had set p to be a small prime, e.g., $p = 31$?

Optional Feedback. Please answer the following questions to help design future problem sets. You are not required to answer these questions, and if you would prefer to answer anonymously, please use the anonymous feedback form. However, we do encourage you to provide feedback on how to improve the course experience.

(a) Roughly how long did you spend on this problem set?

(b) What was your favorite problem on this problem set?

(c) What was your least favorite problem on this problem set?

(d) Any other feedback for this problem set? Was it too easy/difficult?

(e) Any other feedback on the course so far?