WARSAW DOCTORAL SCHOOL OF MATHEMATICS AND COMPUTER SCIENCE

June 18, 2025

Entrance exam

On the following pages you will find 16 problems related to various areas of Mathematics and Computer Science. You are expected to choose and solve **any 4 of them**. Each problem is worth the same number of points.

You are free to choose any problems you wish, i.e. candidates for studies in Mathematics may also choose Computer Science problems, and *vice versa*.

Most of the problems are composed of a few subproblems, but each problem, i.e. all of its subproblems, are graded as a whole.

You may attempt to solve more than 4 problems. All your solutions will be graded, but **only 4 best-graded solutions will contribute to your general grade**.

All your answers should be appropriately justified. **Every problem should be solved on a separate sheet of paper**; of course, the solution of one problem can be written on more that one sheet.

Each sheet of paper should be **signed with your first name and surname**, and **marked with the problem number**.

Exam duration: 3 hours

Good luck!

Analysis

PROBLEM 1. Let γ be the ellipse given by the equation $b^2x^2 + a^2y^2 = a^2b^2$, where $a, b > 0$.

(a) Decide for which $p > 0$ the function $d\colon \mathbb{R}^2 \to [0, \infty)$ defined by $d(\mathbf{x}) = (\operatorname{dist}(\mathbf{x}, \gamma))^p$ is uniformly continuous, where $\operatorname{dist}(\mathbf{x}, \gamma)$ stands for the distance between $\mathbf{x} \in \mathbb{R}^2$ and γ.

(b) Determine a poin on the ellipse γ that is at the shortest distance from the line given by the equation $bx + ay - 2b = 0$.

(c) For $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ let

$$(*) \qquad f(x, y) = \frac{xy}{r^2} - \theta,$$

where $x + \mathrm{i}y = re^{\mathrm{i}\theta}$, $r > 0$, $\theta \in \mathbb{R}$, that is, θ is a (non-uniquely determined) argument of the complex number $x + \mathrm{i}y$. Show that every point $(x, y) \neq (0, 0)$ has a neighborhood U such that for a suitable choice of the argument the function f given by $(*)$ is of class C^1 on U. Calculate the gradient $\nabla f(x, y)$.

(d) Compute the integral of the differential 1-form

$$\oint_\gamma \frac{y^3\,\mathrm{d}x - xy^2\,\mathrm{d}y}{(x^2 + y^2)^2},$$

where the ellipse γ is oriented counterclockwise.

Complex analysis

PROBLEM 2. Let Ω be the complex plane cut along two half-lines $(-\infty, -\frac{1}{2}]$ and $[\frac{1}{2}, \infty)$, i.e. $\Omega = \mathbb{C} \setminus ((-\infty, -\frac{1}{2}] \cup [\frac{1}{2}, \infty))$. In tasks (a) and (b), it suffices to give one example of a map satisfying the desired conditions; however, the answer must be justified.

(a) Determine a Möbius transformation (a homography) that maps the region Ω onto $\mathbb{C} \setminus (-\infty, 0]$.

(b) Determine a one-to-one holomorphic function that maps the region Ω onto the unit disc $\mathbb{D} = \{z \in \mathbb{C}\colon |z| < 1\}$.

(c) Let $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$. Define a holomorphic function $f\colon \mathbb{C} \setminus [\alpha, \beta] \to \mathbb{C}$ by the formula

$$f(z) = \operatorname{Log} \frac{z - \alpha}{z - \beta},$$

where Log stands for the principal branch of the logarithm. Determine the range $f(\mathbb{C} \setminus [\alpha, \beta])$ of the function f.

(d) In the above formula, put $\alpha = -1$, $\beta = 1$, i.e. the function f is given as

$$f(z) = \operatorname{Log} \frac{z + 1}{z - 1}.$$

Let Γ be the positively oriented circle with center at $z_0 = \frac{e+1}{e-1}$ and radius $r = 1$. Compute the integral

$$\int_\Gamma \frac{\mathrm{d}z}{f(z) - 1}.$$

Probability and statistics

PROBLEM 3. Consider the family of horizontal lines

$$\mathcal{L} = \{\{(x, y) \in \mathbb{R}^2 \colon y = k\} \colon k \in \mathbb{Z}\}$$

on the plane $\mathbb{R}^2$. We randomly throw a needle of length 1 onto this plane, with the following distribution. The center of the needle lands inside the square $[0, 10] \times [0, 10]$, with uniform distribution. The direction of the needle has also uniform distribution, and is independent of the position of the center.

(a) Let A be the event that the needle intersects a line from $\mathcal{L}$, and let $X = \mathbb{1}_A$ be the characteristic function of A. Calculate the probability $\mathbb{P}(A) = \mathbb{E}X$.

(b) Let X_n be the number of times the needle intersects a line from $\mathcal{L}$ in n independent tries. Estimate how large n we need to take to know that with probability at least 0.95 we have

$$\left|\frac{1}{n}X_n - \mathbb{E}X\right| \leqslant \frac{1}{10}\mathbb{E}X.$$

(c) What can we say about the asymptotic behavior of

$$\mathbb{P}\left(\frac{1}{n}X_n - \mathbb{E}X\right) > 0.1$$

as $n \to \infty$?

(d) This time we proceed similarly, but now the throws are not independent. The first throw is as above, but in every consecutive throw the direction of the needle must differ from the direction of the previous throw by at least 30° (within the allowed range it still has uniform distribution and is independent from the previous throws and from the position of the center). We also assume that the positions of the center of the needle in all throws are pairwise independent and have, as before, uniform distribution on $[0, 10] \times [0, 10]$. Let Y_n be the number of times the needle intersects a line from $\mathcal{L}$ in n throws. Calculate

$$\lim_{n\to\infty} \frac{\mathbb{E}Y_n}{n}.$$

Geometry and linear algebra

PROBLEM 4. Let $A \in M_{n\times m}(\mathbb{R})$ be an $n \times m$ (not necessarily quadratic) real-valued matrix.

(a) Decide for which A the matrix $B_A = A^T \cdot A \in M_{m\times m}(\mathbb{R})$ defines a scalar product. Are there any pairs $(n, m) \in \mathbb{N} \times \mathbb{N}$ such that for any nonzero matrix A, the matrix B_A defines a scalar product?

(b) Let $f \colon \mathbb{R}^n \to \mathbb{R}^n$ be an isometry. Which complex numbers can be its eigenvalues?

(c) Let $f_\sigma \colon \mathbb{R}^4 \to \mathbb{R}^4$ be an isometry given by a permutation matrix, where $\sigma \in S_4$ is a permutation of four elements. Compute the eigenvalues of f_σ and the Jordan form of $f_\sigma \colon \mathbb{C}^4 \to \mathbb{C}^4$, i.e. when we consider the endomorphism in question over the complex numbers.

(d) Let V be an n-dimensional Euclidean space (that is a vector space with a fixed scalar product). Prove that for any $k \in \{1, \ldots, n\}$ an orthogonal projection does not increase the k-dimensional volume. Start with $k = 1$. Prove that the absolute value of the derminant of a quadratic matrix whose columns have length at most one is at most one.

Algebra

PROBLEM 5. Let D_5 be the group of isometries of regular pentagon. Let A_5 be the group of sign-preserving permutations of five elements.

(a) Find all homomorphisms $h\colon D_5 \times \mathbb{Z}_2 \to A_5$ and $g\colon A_5 \to D_5$.

(b) Let $\mathbb{Z}[i]$ be the ring of Gaussian numbers, that is $\mathbb{Z}[i] = \{x + iy\colon x, y \in \mathbb{Z}\} \subset \mathbb{C}$ with the ring structure inherited from the complex numbers. Prove that $\mathbb{Z}[i]/(3+i)$ is isomorphic to $\mathbb{Z}_m$ for some $m \in \mathbb{N}$. Find m.

(c) Find all possible homomorphisms of rings with unity $\mathbb{Z}[x]/(x^2) \to \mathbb{Z}_{24}$ and $\mathbb{Z}[x, x^{-1}] \to \mathbb{Z}_{24}$.

(d) Prove that a finite domain (a finite ring with no zero divisors) is a field. Find all prime ideals in the ring $\mathbb{Z}_{2025}[x]/(x^{45})$.

Topology

PROBLEM 6. For any cardinal number κ the symbol $D(\kappa)$ stands for the discrete space of cardinality κ. Define $\mathfrak{X} = D(\aleph_0)^{\mathfrak{c}}$, that is, $\mathfrak{X}$ is the product of $\mathfrak{c}$ many copies of the infinite, countable, discrete space.

(a) Construct $\mathfrak{c}$ many pairwise disjoint closed subsets of $\mathfrak{X}$, all of which are homeomorphic copies of the Cantor set.

(b) For each $t \in [0,1]$ define a function $f_t\colon [0,1] \to \{0,1,2,\ldots\}$ by the formulas:

$$f_t(x) = \begin{cases} 0 & \text{if } x = t \\ n & \text{if } n \in \mathbb{N}, \text{ and } \frac{1}{n+1} < |x-t| \leqslant \frac{1}{n}. \end{cases}$$

Define also a function $F\colon [0,1] \to \{0,1,2,\ldots\}^{[0,1]}$ by $F(x)(t) = f_t(x)$ for $x, t \in [0,1]$. Equipping the interval $[0,1]$ with the discrete topology, we may treat the function F as a map from the discrete space $D(\mathfrak{c})$ to the space $\mathfrak{X}$, which is identified with the space of functions $\{0,1,2,\ldots\}^{[0,1]}$. Prove that F is a homeomorphic embedding of $D(\mathfrak{c})$ into $\mathfrak{X}$.

(c) Show that the range $F(D(\mathfrak{c}))$ is a closed subset of $\mathfrak{X}$.

(d) Assume that M is a subset of $\mathfrak{X}$ that is metrizable with the subspace topology inherited from $\mathfrak{X}$. Show that M is nowhere dense, i.e. $\operatorname{int}\operatorname{cl} M = \varnothing$.

Ordinary differential equations

PROBLEM 7. Consider a function $f\colon \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x \ln\left(1 + \dfrac{1}{|x|}\right) & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

(a) Verify whether f satisfies a local Lipschitz condition in a neighborhood of $x = 0$.

(b) Find all solutions of the initial value problem

$$\begin{cases} x'(t) = f(x(t)) \\ x(0) = 0. \end{cases}$$

(c) Show that for every $\varepsilon > 0$ there is a unique forward-time solution $x_\varepsilon \colon [0, +\infty) \to \mathbb{R}$ of

$$\begin{cases} x'(t) = f(x(t)) \\ x(0) = \varepsilon, \end{cases}$$

defined on the whole half-line $[0, +\infty)$. Show also that $\lim\limits_{t \to +\infty} x_\varepsilon(t) = +\infty$.

(d) Examine the stability of the equilibrium point $(x_0, y_0) = (0, 0)$ of the system

$$\begin{cases} x'(t) = f(y(t)) \\ y'(t) = f(x(t)). \end{cases}$$

Functional analysis

Problem 8. For a compact Hausdorff space K we denote by $C(K)$ the Banach space of real-valued continuous functions defined on K, equipped with the supremum norm. Consider the closed subspace $\mathcal{Y}$ of $C([0,1] \times [0,1])$ defined by

$$\mathcal{Y} = \{f \in C([0,1] \times [0,1]) \colon f(x,y) = f(y,x) \text{ for all } 0 \leqslant x, y \leqslant 1\}.$$

By $C([0,1] \times [0,1])/\mathcal{Y}$ we denote the quotient space. For any Banach spaces X and Y we write $X \sim Y$ if there is a linear homeomorphism (an isomorphisms of Banach spaces) from X onto Y. The symbol $X \oplus_1 Y$ stands for the direct sum $X \oplus Y$ equipped with the norm $\|(x,y)\| = \|x\| + \|y\|$.

(a) Show that there exists a closed linear subspace Z of $C([0,1] \times [0,1])$ such that

$$C([0,1] \times [0,1]) \sim C([0,1]) \oplus_1 Z.$$

(b) Decide whether the conjugate space $(C([0,1] \times [0,1])/\mathcal{Y})^*$ contains an isometrically isomorphic copy of the space ℓ_1.

(c) Define a linear operator $T \colon C([0,1] \times [0,1])/\mathcal{Y} \to C([0,1] \times [0,1])$ by the formula

$$T(f + \mathcal{Y})(x,y) = \int_0^1 \int_0^1 e^{tx+uy}\big(f(t,u) - f(u,t)\big)\,\mathrm{d}t\,\mathrm{d}u.$$

Decide whether T is a compact operator.

(d) Let $(a_n)_{n=1}^\infty$ be a sequence of nonnegative numbers. Define a sequence of linear functionals $(\varphi_n)_{n=1}^\infty \subset (C([0,1] \times [0,1])/\mathcal{Y})^*$ by

$$\varphi_n(f + \mathcal{Y}) = \sum_{k=1}^n a_k \int_0^1 \int_0^1 x^k y^k \big(f(x,y) - f(y,x)\big)\,\mathrm{d}x\,\mathrm{d}y.$$

Prove that $(\varphi_n)_{n=1}^\infty$ is pointwise bounded if and only if

$$\sum_{n=1}^\infty \frac{a_n}{n^2} < \infty.$$

Programming languages

PROBLEM 9. The problem consists of four independent subproblems:

(1) Consider the following C code:

```
int f(int *a, int *b){
    int c = *a;
    *a += 6 - *(b + *a) / 2;
    return c;
}

int main(){
    int x = 0;
    int t[5] = {7, 7, 8, 14, 19};
    for (int i = 0; i < 5; ++i)
        t[f(&x, t)] = i + 1;
}
```

What is the content of the array `t` after the `for` loop is executed?

(2) Consider the following Java code:

```
class A {
    int foo(int x){ return x % 2 == 0 ? x / 2 : 3 * x + 1; }
    int bar(int x){ return foo(x / 4); }

    static class B extends A {
        int foo(int x){ return x + 37; }
    }

    static class C extends B {
        int bar(int x){ return super.bar(5);}
    }

    public static void main(String args[]){
        A a = new B();
        B b = new C();
        System.out.println(a.bar(20) + "-" + b.bar(38));
    }
}
```

What will be printed to standard output as a result of executing the `main` method?

(3) Consider the following Haskell code:

```
f [] = 0
f (x:xs) = 1 + f xs

g xs = h xs 0
    where
        h [] a = a
        h (_:xs) a = h xs (a+1)
```

What is the type of the function `h`? What do the functions `f` and `g` do? Which one will terminate faster when both are called with the argument `[1..10^6]`? Justify your answer.

(4) Consider the following function in a language using natural binary encoding:

```
int f(int n) {
    int x = 0;
    while (n > 0) {
        n = n & (n-1);
        x = x + 1;
    }
    return x;
}
```

Propose a potential function that allows you to prove that execution of the function `f` always terminates. What does the function `f` return when called with a positive `n`?

Discrete mathematics

PROBLEM 10. Consider the domain of simple (i.e., without loops or multi-edges), undirected graphs. Let n denote the number of vertices in the graph. Each answer should present a proof or a constructive example of a graph.

(a) Does there exist a bipartite graph for which $n = 2025$ that contains an Eulerian cycle?

(b) Let the vertices of a graph be numbered with natural numbers from 0 to $n-1$, and let two vertices i, j be connected by an edge if and only if $|i-j| \leqslant 2$. How many distinct paths are there from vertex 0 to vertex $n-1$ such that we always move to a vertex with a greater number?

(c) Let the set of vertices consist of all strings of length k over the alphabet $\{A, B, C\}$, and let two vertices be connected by an edge if and only if their strings differ on exactly one position. What is the (vertex) chromatic number of this graph for a given k?

(d) Let d be the maximum degree of a vertex in the graph. Prove that every connected graph has a matching of size at least $\lfloor \frac{n}{2d} \rfloor$.

Algorithms and data structures

PROBLEM 11. Let s be a non-empty string over the alphabet $\Sigma = \{\texttt{0}, \texttt{1}, \ldots \texttt{9}\} \cup \{\texttt{|}\}$, consisting of digits and the vertical bar symbol. We call s *valid* if it starts and ends with a digit and no two consecutive symbols in s are vertical bars. A valid string s can be interpreted as a sequence of numbers separated by vertical bars, e.g. $s = \texttt{18|06|2025}$ corresponds to the sequence of numbers $18, 6, 2025$. Let the *sum* of s, denoted $\mathbf{sum}(s)$, be the sum of these numbers, so for s above we have $\mathbf{sum}(s) = 18 + 6 + 2025 = 2049$. In this problem, we are interested in computing the sum of a dynamically changing valid string s. We consider data structures supporting the following operations:

init(v,p): Set s to v, which is guaranteed to be a valid string. Also, the structure does not need to report the actual sums (which are rather large) but rather their reminders modulo a prime number $p > 5$. This operation is guaranteed to be used exactly once, as the first operation.

add(i): Add a vertical bar in front of the i-th digit in s. It is guaranteed that there is no bar there prior to this operation.

remove(i): Remove the vertical bar in front of the i-th digit in s. It is guaranteed that it exists.

sum(): Return the sum of s modulo p.

Your task is as follows:

(a) Design an efficient data structure supporting these operations.

(b) Design an efficient data structure for the case where the **add** operation is never used.

(c) Consider the case where the initial value of the string s contains no vertical bars and the total number of operations performed k satisfies $k = o(n)$. Design a data structure for this case such that the total time complexity of these k operations is of the form $O(n + f(k))$.

(d) Like in previous case we consider $k = o(n)$ and look for a solution with time complexity $O(n + f(k))$. However, this time the initial value of the string s is a valid prefix of the infinite string `1|2|3|1|2|3|`...

Remark. You need to justify the correctness of your algorithms and estimate their time complexity. The score in all subproblems depends very strongly on the time complexity of your solution.

Logic and databases

Problem 12. In this problem, we consider undirected graphs without loops. In other words, the relation $E \subseteq V^2$ is always irreflexive and symmetric.

(1) Prove that there is no formula ϕ of first order logic, over the signature consisting of the single binary relation symbol E, such that a finite graph (V, E) is Eulerian if and only if the structure (V, E) satisfies ϕ.

(2) Does there exist a formula ϕ of first order logic, over the signature consisting of the binary relation symbol E and additional relational symbols, such that the finite graph (V, E) is Eulerian if and only if there is an interpretation of the other relational symbols for which the obtained structure $(V, E \ldots)$ satisfies ϕ?

(3) Does there exist a formula ϕ of first order logic, over the signature consisting of the single binary relation symbol E, which is satisfied for infinitely many finite graphs, but no infinite graph?

(4) Does there exist a formula ϕ of first order logic, over the signature consisting of the single binary relation symbol E, which is satisfied for some infinite graph, but no finite graph?

Automata and formal languages

Problem 13. Consider a set X with operation $\ominus : X \times X \to X$, a finite subset $A \subseteq X$, and an element $z \in X$. Let $L \subseteq A^*$ be the language of words $x_1 \ldots x_n$ such that $n \geqslant 1$ and it is possible to add parentheses to

$$x_1 \ominus x_2 \ominus x_3 \ominus \ldots \ominus x_n$$

in such a way that z is the result of the obtained expression. For example, for $X = \mathbb{Z}$, $x \ominus y = x - y$, $z = 0$, $A = \{2, 4, 6\}$, the word 462 is in L, since $4 \ominus (6 \ominus 2) = 0$. In the following cases, determine whether L must be a regular language, and whether it must be a context-free language. Explain your answer.

(1) $(X, \ominus, z)$ is a finite group

(2) $(X, \ominus, z)$ is an Abelian group

(3) X is a finite set

(4) $X = \mathbb{Z}$, $x \ominus y = x - y$ (A is any finite subset of X and z is any element of X)

Computation theory and computational complexity

PROBLEM 14. For each of the given problems decide whether it is **NP**-complete (assuming **P**≠**NP**). Justify your answers.

(a) Given is a family $\mathcal{F}$ of 2-element subsets of the universe $\mathcal{U}$, and an integer $k > 0$. Decide whether there exists a subfamily $\mathcal{F}' \subseteq \mathcal{F}$ of size k, and such that the elements of $\mathcal{F}'$ cover $\mathcal{U}$.

(b) Given is an n-element multi-set $\mathcal{X}$ of natural numbers. Decide whether $\mathcal{X}$ can be partitioned into $\frac{n}{2}$ multi-sets with equal sums.

(c) Given is an undirected bipartite graph G and a non-negative integer k. Decide whether there exists a set of vertices I in G of size k, and such that no two vertices of I are connected by an edge of G.

(d) Given is a non-negative integer k, and an undirected graph G, in which all vertices have degrees at least k. Decide whether G contains a clique of size k.

Concurrent and distributed programming, computer systems

PROBLEM 15. *Concurrent data structures* may be accessed and modified by many threads without the need for additional synchronization. An implementation of a concurrent data structure provides methods that can be invoked concurrently by many threads: the implementation guarantees correctness by careful synchronization. Consider a *concurrent (max) heap* that stores up to `N` items of type `T`. The heap provides two methods: `void insert(T item)` that inserts an element; and `T extract()` that removes and returns the maximal item from the previously-inserted items (assume `T` implements comparators `<` and `>`). Both methods block the calling thread until successful completion. Write the pseudo-code for the following, using primitive datatypes (e.g., scalars, arrays) and *semaphores* (use the standard, weakly-fair semaphore sematics). Comment on the efficiency and semantics of your implementations.

(1) Implement basic variants, `basicInsert` and `basicExtract` that are correct, but not necessarily efficient.
(2) Implement an efficient (and correct) version of `insert`, `concInsert`. Assume that `extract` is not called in parallel to any invocation of `concInsert` and that it is implemented correctly.
(3) Now assume you are given implementations `concInsert` and `concExtract`. Each of these correctly handles many concurrent invocations. However, a call by a thread to `concExtract` while any other thread executes `concInsert` corrupts the data structure (and vice versa). Using calls to `concInsert` and `concExtract`, implement efficient and correct `insert` and `extract`.

Bioinformatics

PROBLEM 16. In the human genome, the `CpG` dinucleotide (cytosine followed by guanine) occurs relatively rarely. This is because C in the `CpG` context is often methylated (converted to 5-methylcytosine), and 5-methylcytosine is prone to deamination to thymine, leading to $C \to T$ mutations. Near the beginnings of genes (in promoter regions), methylation occurs less frequently, which favors the preservation of `CpG` dinucleotides and the formation of so-called "`CpG` islands." Therefore, the human genome can be divided into:

- regions rich in `CpG` (`CpG` islands), and
- background regions (poor in `CpG`).

We model this division using a hidden Markov model (HMM) with eight hidden states:

- A^+, C^+, G^+, T^+ — states corresponding to nucleotides within a `CpG` island,
- A^-, C^-, G^-, T^- — states corresponding to nucleotides outside a `CpG` island.

Emissions are deterministic, i.e., for example, the state G^+ always emits G, and C^- always emits C, etc.

Transition probabilities:

$P(C^+ \to G^+) = 0.4$,
$P(C^- \to G^-) = 0.05$,
All other transitions within the same group (e.g., $A^+ \to T^+, G^- \to C^-,$ etc.) $= 0.25$,
Transitions between regions (e.g., $A^+ \to C^-, G^- \to T^+$) $= 0.01$.

The transition matrix is normalized in each row (the probabilities sum to 1 for each state).

Initial probabilities:

$$P(\text{start in state } +) = P(\text{start in state } -) = 0.5.$$

(1) For the sequence $A\,C\,G\,C\,G$, propose two different state paths: one assuming the entire sequence comes from a CpG region, and the other assuming it comes from the background. Calculate their total probabilities in the model. Based on the calculated probabilities, indicate which hypothesis (CpG vs. non-CpG) is more likely for this sequence.
(2) The occurrence of CpG islands and background sequences could be modeled using two states instead of eight. In your opinion, does using eight states offer any advantage? Justify your answer.
(3) Estimate the most probable hidden state path for the sequence $A\,C\,G\,C\,G$ using the Viterbi algorithm. The path does not have to come entirely from one type of region.