Вступительное испытание НИУ ВШЭ 2019 г. Демонстрационный вариант по направлению 01.04.02 Прикладная математика и информатика, ОП "Науки о данных"

Время выполнения задания — 240 мин.

Решения заданий должны быть записаны по-русски или по-английски. Каждая задача оценивается из 10 баллов, максимальная сумма — 100 баллов. Допускается использование бухгалтерского калькулятора и таблиц статистических распределений; черновики собираются, но не проверяются.

Задача 1 (Теория чисел). При каких натуральных n число $5^{2n+1}2^{n+2}+3^{n+2}2^{2n+1}$ делится нацело на 19?

Задача 2 (Математический анализ). Вычислить двойной интеграл $\int \int_D (x+y) dx dy$, если область D ограничена кривой $x^2 + y^2 = x + y$.

Задача 3 (Алгебра). Пусть матрица A размера 3×3 такова, что для любого вектора столбца $v \in R^3$ вектора Av и v ортогональны. Доказать, что $A^t + A = 0$, где A^t транспонированная матрица.

Задача 4 (Алгоритмы). Дано дерево на n вершинах. На рёбрах дерева написаны целые числа в двоичной записи (можее считать, что все имеют одинаковое количество двоичных разрядов).

Требуется уметь находить xor всех чисел на рёбрах пути из вершины u в вершину v для любых заданных пользователем вершин u в вершину v. Операция xor — это побитовое исключающее или. Например, 1010 xor 1100 = 0110.

- (а) Во многих ситуациях бывает полезно сделать *предподсчёт*: заранее произвести некоторые вычисления и сохранить их результаты в памяти, чтобы в дальнейшем ускорить вычисления. Простой пример: если нам дан некоторый текст и нужно уметь быстро отвечать на запросы вида "сколько раз слово х встречается в тексте после слова у", то можно заранее подсчитать, сколько раз встречается в тексте каждое словосочетание из двух слов, и это позволит в дальнейшем очень быстро отвечать на вопрос (за стоимость запроса к памяти). Придумайте какой-либо вариант предподсчёта для задачи вычисления хог вдоль путей. Оцените, сколько операций потребует предпосчёт, сравните число операций, требуемое для вычисления хог вдоль пути без предподсчёта и с использованием предложенного вами способа предподсчёта.
- (b) Придумайте, как выполнить некоторый предпосчёт за время O(n), который позволит впоследствии отвечать на запрос **xor** за время O(1).

Задача 5 (Алгоритмы). На плоскости задан набор точек $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

- (a) Предложите алгоритм, определяющий, обладает ли множество этих точек осью симметрии, параллельной оси у. Оцените количество арифметических операций и операций сравнения, а также количество дополнительной памяти, которое будет использовать ваш алгоритм.
 - (b) Оптимизируйте алгоритм, чтобы он работал за время $O(n \log n)$.

Задача 6 (Графы). В плоском двусвязном графе 7 граней (считая внешнюю) – 3 треугольника, 3 четырёхугольника и пятиугольник. Сколько в этом графе вершин и рёбер? Приведите пример диаграммы графа.

Задача 7 (Оптимизация). Решить задачу оптимизации:

$$min_{x\in R^n}\sum_{i=1}^n x_i(\ln x_i-c_i)$$
 при условии, что $\sum_{i=1}^n x_i=1$, где

 $c \in \mathbb{R}^n_+$ – заданный вектор с неотрицательными элементами.

Задача 8 (Алгебра). Матрицей перестановки порядка n называется такая матрица размера n на n, составленная из 0 и 1, что сумма (в поле действительных чисел) элементов по каждому ее столбцу и каждой строке равна 1. Пусть $\lambda_1, \lambda_2, \ldots, \lambda_5$ — собственные числа матрицы перестановки порядка 5

$$P_{\pi}^{\{5\}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Найдите $\lambda^* = \min_i |\lambda_i|$

Задача 9 (Теория вероятностей). Независимые одинаково распределенные случайные величины $X_1, X_2, \dots X_{20}$ принимают только значения 2 и 3, при этом значение 3 принимается с вероятностью 0,2. Найти вероятность того, что сумма данных величин будет равна 46.

Задача 10 (Математическая статистика). Программист желает найти доверительный интервал для математического ожидания времени работы программы с надежностью $\gamma=0,95$. Известно, что выборочное среднее времени работы 65 циклов программы составило 0,09 сек. Найдите искомый доверительный интервал в предположении, что времена работы программы за различные циклы независимы и распределены нормально с постоянными параметрами, при этом генеральное среднеквадратичное отклонение равно 0,02.