

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**Директор физтех-школы
прикладной математики и
информатики**

А.М. Райгородский

Программа государственной итоговой аттестации

**Подготовка к сдаче и сдача государственного экзамена по информатике и вычислительной
технике**

по направлению:	Информатика и вычислительная техника
профиль подготовки:	Прикладная математика, компьютерные науки и инженерия Физтех-школа Прикладной Математики и Информатики
курс:	4
квалификация:	бакалавр
семестр:	7 (Осенний)

Программу составили:

В.В. Яковлев, канд. физ.-мат. наук, заведующий кафедрой

К.Ю. Войтиков, канд. техн. наук, доцент, доцент

Программа обсуждена на заседании кафедры алгоритмов и технологий программирования 04.06.2020

1. Цели и задачи

Цели

Целью государственного экзамена по математике является установление уровня подготовки обучающегося по математическим дисциплинам и соответствия результатов освоения обучающимся образовательной программы требованиям образовательного стандарта по направлению подготовки.

Задачи

- * оценка степени освоения обучающимися теоретических положений основных дисциплин, формирующих специальные знания в рамках освоения образовательной программы;
- * оценка умения применять полученные знания для решения конкретных задач;
- * оценка актуальности полученных студентами знаний и их соответствие требованиям потенциальных работодателей.

2. Перечень компетенций, уровень сформированности которых оценивается при проведении государственного экзамена

Код и наименование компетенции	Индикаторы достижения компетенции
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.4 Грамотно, логично, аргументированно формирует собственные суждения и оценки
ПК-1 Способен ставить, формализовывать и решать задачи, в том числе разрабатывать и исследовать математические модели изучаемых явлений и процессов, системно анализировать научные проблемы, получать новые научные результаты	ПК-1.3 Способен применять теоретические и (или) экспериментальные методы исследований к конкретной научной задаче и интерпретировать полученные результаты
ПК-2 Способен самостоятельно или в качестве члена (руководителя) малого коллектива организовывать и проводить научные исследования и их апробацию	ПК-2.3 Способен проводить апробацию результатов научно-исследовательской работы посредством публикации научных статей и участия в конференциях

3. Перечень примерных вопросов, выносимых на государственный экзамен

Информатика

1. Динамический массив. Амортизационный анализ. Учетная оценка времени добавления элемента в динамический массив (с удвоением заполненного буфера).
2. Связные списки. Стек, очередь, дек и их реализации.
3. Быстрая сортировка (QuickSort). Поиск порядковой статистики методом “Разделяй и властвуй” (QuickSelect).
4. Сортировка слиянием (MergeSort). Поразрядные сортировки.
5. Двоичная куча и сортировка кучей (HeapSort). Слияние k отсортированных массивов с помощью кучи.
6. Хеш-таблица, полиномиальная хэш-функция.
7. Динамическое программирование: общая идея, линейная динамика, матричная, динамика на отрезках.
8. RMQ. Sparse table. Дерево отрезков.
9. LCA: сведение к RMQ и метод двоичного подъёма.
10. Двоичное дерево поиска. Обходы в глубину и в ширину. Поиск ключа, наивные вставка и удаление ключа. АВЛ-дерево. Красно-чёрное дерево.

- 11 . Декартово дерево. Декартово дерево по неявному ключу.
12. Минимальное остовное дерево: алгоритмы Прима и Крускала.
- 13 . Максимальные потоки в сети. Методы: Форда-Фалкерсона; Эдмондса-Карпа (б/д).
- 14 . Обход графа в глубину, ширину.
- 15 . Поиск кратчайших путей в графе: алгоритмы Дейкстры, Форда-Беллмана, Флойда-Уоршелла.
- 16 . Поиск сильно-связных компонент в графе.
- 17 . Мосты и точки сочленения в графе.
- 18 . Нахождение подстроки в строке: префикс-функция, алгоритм Кнута-Морриса-Пракка.
- 19 . Стандартные контейнеры: `vector`, `deque`, `queue`, `priority_queue`, `set`, `map`, итераторы, компараторы.
- 20 . Бор. Алгоритм Ахо-Корасика.
- 21 . Суффиксное дерево.
- 22 . Вычисление выпуклой оболочки множества точек в 2D и 3D.
- 23 . Триангуляция Делоне. Диаграмма Вороного.
- 24 . Длинная арифметика. Сложение, вычитание, хранение знака, умножение, деление. Алгоритм Карацубы.
25. Постановка задачи обучения с учителем (`supervised learning`), постановка задачи обучения без учителя (`unsupervised learning`). Постановка задачи классификации и метрики качества классификации. Наивный Байесовский классификатор.
26. Постановка задачи регрессии и метрики качества регрессии. Линейная регрессия. Теорема Гаусса-Маркова (формулировка). Проблема мультиколлинеарных признаков. $L1$ и $L2$ регуляризация, их влияние на веса признаков.
27. Проблема несбалансированных классов. Работа с категориальными признаками и пропущенными значениями. `Mean encoding`. Примеры простых алгоритмов, решающих стандартные задачи: `kNN`, наивный байесовский классификатор.
28. Логистическая регрессия. Понятия отступа (`Margin`). Эквивалентность решений полученных методом максимального правдоподобия и минимизации логистической функции потерь. Логистическая функция потерь, кросс-энтропия.
29. Решение задачи классификации методом опорных векторов. Понятия отступа (`Margin`). `Hinge-loss`. Оптимизационная задача в `SVM` (декларативно).
30. Процедура построения решающего дерева, критерии информативности: энтропийный, Джини. Бустинг (принцип построения ансамбля).
31. `Bias-Variance decomposition` (декларативно). Процедура `bootstrap`, алгоритм `bagging`. Метод случайных подпространств (`RSM`), `Random Forest`.
32. Градиентный бустинг. Принцип построения. Какие алгоритмы могут использоваться в качестве базовых алгоритмов.
33. Задача снижения размерности: алгоритмы `PCA` и `t-SNE`. Связь `PCA` и `SVD`. Может ли `PCA` давать не единственное решение?
34. Проблема переобучения. Пример, причины возникновения (идейно). Кросс-валидация. Мотивация использования `train`, `val` и `test` выборок. Понятие параметров и гиперпараметров моделей. Процедура подбора гиперпараметров. Понятие регуляризации (в общем случае). Способы регуляризации различных моделей (линейные модели, деревья, ансамбли, нейронные сети).
35. Метод градиентного спуска. Метод обратного распространения ошибки (`backpropagation`). Функции активации (`Sigmoid`, `tanh`, `ReLU`), их свойства и проблемы. Функции потерь в задаче многоклассовой классификации и регрессии. Методы регуляризации нейронных сетей: `Dropout`, `Batch normalization`, `data augmentation`.
36. Методы регуляризации нейронных сетей: `Dropout`, `Batch normalization`, `data augmentation`. Слабые стороны стохастического градиентного спуска. Способы доработки: `Momentum`, `Nesterov momentum`, `RMSprop`, `Adam`. Минусы данных подходов.
37. Рекуррентные нейронные сети (`RNN`) для упорядоченных данных. Основные принципы работы. Проблема затухающего градиента и ее возможные решения. Рекуррентные блоки: наивный (`Vanilla RNN`), `LSTM`, `GRU`, мотивация их использования. Функции активации в рекуррентных блоках.

38. Методы работы с изображениями. Почему линейные слои не получили широкого применения в задаче компьютерного зрения? Сверточные слои в нейронной сети. Мотивация их использования в задачах анализа изображений и сигналов. Одномерные и двумерные свертки (Conv1d и Conv2d). Max & average pooling.
39. Операционные системы и их компоненты. Ядро операционных систем. Системные вызовы и их отличия от обычных библиотечных функций. Способы реализации системных вызовов (прерывания, sysenter, syscall).
40. Целочисленная арифметика в представлении компьютера. Знаковые и беззнаковые значения, способы представления отрицательных значений. Целочисленное переполнение и его контроль. Длинная целочисленная арифметика.
41. Вещественная арифметика. Представления с фиксированной и плавающей точкой. Стандарт IEEE754. Специальные вещественные значения, определенные стандартом IEEE754 и операции над ними.
42. Процессы и потоки. Сходства и различия между ними. Реализация многозадачности и алгоритмы планирование задач в операционных системах.
43. Проблема многопоточной синхронизации. Атомарные переменные и объекты блокировки. Неблокирующие структуры данных и их реализация.
44. Message Passing Interface (MPI) Существующие реализации, задачи MPI как среды программирования. Жизненный цикл MPI программы. Создание и завершение MPI процессов. Организация потока ввода-вывода, параметры указываемые MPI программе.
45. Понятие ускорения и масштабируемости параллельных программ. Закон Амдала. Оценка эффективности параллельных программ. Ярусно-параллельная форма программы.
46. Распределенные файловые системы. Роли элементов системы, обеспечение отказоустойчивости. Алгоритмы чтения и записи в распределенных файловых системах. Репликация данных.
47. Модель вычислений MapReduce. Пары ключ-значение в реализациях MapReduce. Основные стадии вычислений и дополнительные элементы модели.
48. Соединение данных (операция Join) в модели MapReduce. Модель вычислений и оптимизации данных.
49. Итеративные вычисления на больших объемах данных. Модель ленивых вычислений и структура хранения данных в реализации Spark (RDD). Кэширование результатов вычислений и итеративные вычисления.
50. Распределенные диспетчеры сообщений. Репликация и реализация отказоустойчивости. Семантики доставки сообщений.
51. Степени изоляции транзакций. Принципы атомарности, согласованности, изолированности и устойчивости (правила ACID). Применение (commit) и отмена транзакций.
52. Теорема Фишера-Линч-Патерсона (FLP-теорема), CAP-теорема и их применение. Распределенные системы хранения конфигураций.
53. Выбор процесса-лидера или машины-лидера в распределенных системах. Алгоритм консенсуса.

4. Порядок сдачи государственного экзамена

К государственному экзамену допускается обучающийся, освоивших дисциплины, который покрываются программой экзамена, и не имеющий по ним академических задолженностей.

Государственный экзамен состоит из двух частей: 1) по дисциплинам из области математики (разделы 1-3 программы экзамена); по дисциплинам из области алгоритмов и компьютерных технологий (разделы 4-7 программы экзамена). На каждую часть экзамена отводится отдельный день проведения.

Формат экзамена - устный.

1. Порядок сдачи части по дискретной математике

Экзаменационный билет состоит из двух вопросов, выбор которых осуществляется учащимися из разделов [1-3], по одному вопросу из двух разных разделов. Опрос студентов о выборе вариативных разделов проводится не менее чем за одну неделю до даты государственного экзамена

Экзаменационные билеты разрезаны на несколько частей, сгруппированных по разделам экзамена, студент вытягивает разные части билета из соответствующих стопок. На подготовку к устному ответу студенту отводится 1 астрономический час. Во время подготовки к ответу разрешается пользоваться «официальной шпаргалкой», материалами курсов.

Устная часть экзамена включает в себя ответ студента на вопросы экзаменационного билета. После завершения устного ответа члены ГЭК могут задавать дополнительные и уточняющие вопросы, в том числе требующие дополнительного времени на подготовку.

2. Порядок сдачи части по алгоритмам и компьютерным технологиям

Экзаменационные билеты состоят из трех вопросов: одного из обязательного раздела [4], и одного вопроса - из разделов [5 или 6] на усмотрение студента. Опрос студентов о выборе вариативных разделов проводится не менее чем за одну неделю до даты государственного экзамена.

Экзаменационные билеты разрезаны на несколько частей, сгруппированных по разделам экзамена, студент вытягивает разные части билета из соответствующих стопок. На подготовку к устному ответу студенту отводится 1,5 астрономических часа. Во время подготовки к ответу разрешается пользоваться заранее приготовленными рукописными записями без ограничения по объему.

Устная часть экзамена включает в себя ответ студента на вопросы экзаменационного билета. После завершения устного ответа члены ГЭК могут задавать дополнительные и уточняющие вопросы, в том числе требующие дополнительного времени на подготовку.

3. Особенности сдачи экзамена в условиях ограничений, связанных с COVID-19

В случае действия ограничительных мер, запрещающих проведение экзамена в присутственной форме, экзамен проводится с применением дистанционных средств, обеспечивающих возможность видеосвязи.

Перед началом экзамена и получением билета студент обязан предоставить (показать на камеру) документ, удостоверяющий личность: зачетную книжку, студенческий билет, или паспорт.

Во время подготовки к ответу разрешается пользоваться любыми источниками информации, включая электронные, но время на подготовку сокращается до 30 минут.

После завершения устного ответа члены ГЭК обязаны задать по одному дополнительному вопросу из каждого раздела программы экзамена, к которым готовился студент. Дополнительные вопросы не подразумевают продолжительное время на подготовку, и требует краткий ответ.

5. Описание материально-технической базы, необходимой для проведения государственного экзамена

Аудитория для проведения аттестационного испытания, оснащенная местами для обучающихся и экзаменационной комиссии.

В случае действия ограничительных мер, запрещающих проведение экзамена в присутственной форме - работоспособные дистанционные рабочие места для всех обучающихся и экзаменаторов.

6. Перечень рекомендуемой литературы

Основная литература

1. Программирование: теоремы и задачи [Текст], [учеб. пособие] /А. Шень. -М., МЦНМО, 2011
2. Компьютерные сети: принципы, технологии, протоколы, учебное пособие для вузов : рек. М-вом образования и науки РФ /В. Г. Олифер, Н. А. Олифер. Санкт-Петербург, Питер, 2019
3. Прикладная статистика. Принципы и примеры [Текст] : [учеб. пособие для вузов] / Д. Кокс, Э. Снелл ; пер. с англ. Е. В. Чепурина ; под ред. Ю. К. Беляева .— М. : Мир, 1984 .— 200 с.
4. Архитектура компьютера, [учеб. пособие для вузов] / Э. Таненбаум, Т. Остин. — Санкт-Петербург, Питер, 2020.— URL: <https://ibooks.ru/bookshelf/361850/reading> (дата обращения: 26.11.2020). - Полный текст (Режим доступа : из сети МФТИ / Удаленный доступ)

Дополнительная литература

1. Алгоритмы: построение и анализ [Текст], [учебник для вузов] /Т. Кормен [и др.] ; [пер. с англ. И. В. Красикова и др.]. Санкт-Петербург, Диалектика, 2019

7. Рекомендации обучающимся по подготовке к государственному экзамену

При подготовке к устной части государственного экзамена обучающимся рекомендуется вспомнить темы математических дисциплин, входящие в программу устной части государственного экзамена, используя при необходимости конспекты лекций и рекомендуемую литературу. После повторения каждой темы обучающемуся рекомендуется самостоятельно написать формулировки и доказательства теорем, содержащихся в программе устной части государственного экзамена, без использования литературы и вспомогательных средств.

Если возникают вопросы, которые студент не может самостоятельно решить с помощью рекомендуемой литературы, эти вопросы рекомендуется задать на консультации, проводимой преподавателем кафедры по соответствующей дисциплине.

8. Методика и критерии оценки государственного экзамена

За каждую из двух частей экзамена выставляется отдельная оценка по 10-балльной шкале. Оценки ниже 3 считаются неудовлетворительными и блокирующими: в случае получения неудовлетворительной оценки за любую часть экзамена, экзамен считается сданным на оценку "неудовлетворительно", и до сдачи другой части экзамена обучающийся не допускается.

Ответы на устную часть по каждому из разделов программы оцениваются отдельно по 10-балльной шкале. Оценка за часть экзамена, которая проводится в один день, определяется как среднее арифметическое из оценок за различные разделы. Результат оценки за часть округляется до целого значения по стандартным арифметическим правилам и объявляется обучающемуся в день проведения части экзамена.

Итоговая оценка выставляется по 10-балльной шкале и определяется как среднее значение из двух оценок, полученных за разные части экзамена. Способ округления - стандартный арифметический.

отлично (10) – правильный, четкий и уверенный ответ на оба вопроса билета и дополнительные вопросы;

отлично (9) – даны правильные ответы на оба вопроса билета и дополнительные вопросы с незначительными неточностями;

отлично (8) – даны ответы на оба вопроса билета и дополнительные вопросы после небольших исправлений и наводящих вопросов экзаменаторов;

хорошо (7) – даны ответы на оба вопроса билета, но нет верного ответа на один из дополнительных вопросов;

хорошо (6) – есть недочеты в ответе на один из вопросов билета и нет верного ответа на один из дополнительных вопросов;

хорошо (5) – есть недочеты в ответах на оба вопроса билета и нет верного ответа на один из дополнительных вопросов;

удовлетворительно (4) – есть недочеты в ответах на оба вопроса билета или нет ответа ни на один из дополнительных вопросов;

удовлетворительно (3) – нет ответа на один из вопросов билета, но есть ответы на дополнительные вопросы (возможно с недочетами);

неудовлетворительно (2) – нет ответа на один из вопросов билета и на дополнительные вопросы;

неудовлетворительно (1) – нет ответа ни на один из вопросов билета.

9. Особенности проведения государственной итоговой аттестации для инвалидов и лиц с ограниченными возможностями здоровья

Для обучающихся из числа инвалидов государственная итоговая аттестация проводится с учетом особенностей их психофизического развития, их индивидуальных возможностей и состояния здоровья (далее – индивидуальные особенности).

10.1. При проведении ГИА обеспечивается соблюдение следующих общих требований:

– проведение государственной итоговой аттестации для инвалидов в одной аудитории совместно с обучающимися, не имеющими ограниченных возможностей здоровья, если это не создает трудностей для обучающихся при прохождении ГИА;

– присутствие в аудитории ассистента (ассистентов), оказывающего обучающимся инвалидам необходимую техническую помощь с учетом их индивидуальных особенностей (занять рабочее место, передвигаться, прочитать и оформить задание, общаться с членами ГЭК);

– пользование необходимыми обучающимся инвалидам техническими средствами при прохождении ГИА с учетом их индивидуальных особенностей;

– обеспечение возможности беспрепятственного доступа обучающихся инвалидов в аудитории, туалетные и другие помещения, а также их пребывания в указанных помещениях.

10.2. По письменному заявлению обучающегося инвалида продолжительность выступления обучающегося при защите выпускной квалификационной работы – не более чем на 15 минут.

10.3. Обучающийся инвалид не позднее, чем за 3 месяца до начала проведения ГИА подает письменное заявление о необходимости создания для него специальных условий при проведении государственных аттестационных испытаний с указанием особенностей его психофизического развития, индивидуальных возможностей и состояния здоровья. К заявлению прилагаются документы, подтверждающие наличие у обучающегося индивидуальных особенностей (при отсутствии указанных документов в дирекции института).

В заявлении обучающийся указывает на необходимость (отсутствие необходимости) присутствия ассистента на государственном аттестационном испытании, необходимость (отсутствие необходимости) увеличения продолжительности выступления при защите выпускной квалификационной работы по отношению к установленной продолжительности.

10. Примеры контрольных заданий, билетов

Примеры заданий приведены в приложении

10. Примеры контрольных заданий, билетов

10.1. Примеры билетов на часть экзамена по разделам [1-3]

Билет №1

1. Понятия множества и подмножества. Операции над множествами, тождества. Отображения и соответствия. Сравнение множеств по мощности. Теорема Кантора–Бернштейна. Счётные множества и их свойства. Теорема Кантора.
2. Хроматическое число, число независимости, кликовое число. Нижняя оценка хроматического числа через число независимости и через кликовое число; сравнение порядка этих оценок для случайного графа $G(n, 1/2)$ в модели Эрдёша–Реньи (а.п.н. $\alpha(G) \leq 2 \log_2(n)$). Теорема Эрдёша о существовании графов с произвольно большим обхватом и хроматическим числом.

Билет № 2

1. Числа Рамсея: определение, и точные значения $R(s, t)$ при $s \leq 3, t \leq 4$. Верхняя оценка Эрдёша–Секереша, её следствие для диагональных чисел Рамсея; нижняя оценка диагональных чисел с помощью простого вероятностного метода.
2. ходимость случайных величин: по вероятности, по распределению, почти наверное, в среднем. Связь между сходимостями (б/д). Лемма Слуцкого (б/д). Теорема о наследовании сходимости. Дельта-метод.

Билет №3

1. Доверительные интервалы. Метод центральной статистики. Метод построения асимптотических доверительных интервалов.
2. Формальная арифметика. Примеры выводов в аксиоматике Пеано. Моделирование машин Тьюринга в формальной арифметике (б/д). Теорема Гёделя о неполноте: формулировка и идея доказательства.

10.2. Примеры билетов на часть экзамена по разделам [4-6]

Билет №1

Обход графа в глубину, ширину.

Проблема переобучения. Пример, причины возникновения (идейно). Кросс-валидация. Мотивация использования train, val и test выборок. Понятие параметров и гиперпараметров моделей. Процедура подбора гиперпараметров. Понятие регуляризации (в общем случае). Способы регуляризации различных моделей (линейные модели, деревья, ансамбли, нейронные сети).

Билет №2

Минимальное остовное дерево: алгоритмы Прима и Крускала.

Решение задачи классификации методом опорных векторов. Понятия отступа (Margin). Hinge-loss. Оптимизационная задача в SVM (декларативно).

Билет №3

1. LCA: сведение к RMQ и метод двоичного подъёма.

2. Задача снижения размерности: алгоритмы PCA и t-SNE. Связь PCA и SVD. Может ли PCA давать не единственное решение?