

**«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)**

Председатель приемной комиссии

Д.А. Ендовицкий

21 октября 2022



09.04.04 Программная инженерия

2022

Программа разработана на основе ФГОС высшего образования по программе бакалавриата 09.03.04 Программная инженерия.

1. Аннотация к программе по направлению подготовки Программная инженерия (очная форма обучения)

«Системное программирование»

Руководитель магистерской программы – заведующий кафедрой программирования и информационных технологий, доктор физ.-мат. наук, доцент Махортов С.Д.

Краткое описание магистерской программы:

Программа посвящена изучению основных категориальных понятий и технологий разработки программных систем, системного и конструктивного подходов, проблем декомпозиции, понятий организационной и функциональной структуры информационной системы, этапов микро- и макро-проектирования и их характеристик, форм представления информационных систем, методик конструирования моделей данных и знаний, субъективной производительности информационной системы, критериев эффективности программного обеспечения. Объектами профессиональной деятельности магистра являются научные исследования; проектирование, разработка и тестирование информационных систем и программного обеспечения, в том числе системного уровня; руководство научно-исследовательскими и опытно-конструкторскими разработками в области информатики и вычислительной техники.

Вступительное испытание – по дисциплине «Компьютерные и математические науки». Форма вступительного испытания: письменный экзамен.

Программа вступительного испытания по дисциплине «Компьютерные и математические науки»

2. Составители: Крыловецкий А.А., декан ФКН, Борисов Д.Н., доцент, зав. кафедрой ИС, Сирота А.А., профессор, зав. кафедрой ТООИ, Махортов С.Д., доцент, зав. кафедрой ПиИТ, Кургалин С.Д., профессор, зав. кафедрой ЦТ, заместитель декана ФКН по учебной работе, Атанов А.В. доцент кафедры ЦТ, доцент, заместитель декана ФКН по учебной работе

3. Основные знания, умения и навыки, которыми должен обладать поступающий.

Должен знать:

- основные понятия и методы математического анализа, линейной алгебры, элементов математической логики, дискретной математики, теории дифференциальных уравнений и элементов теории уравнений математической физики, теории вероятностей и математической статистики, случайных процессов, статистического оценивания и проверки гипотез, технологий обработки информации, элементов теории функций комплексного переменного.

- операционные системы, принципы управления сетью ЭВМ; один из языков программирования, структуру локальных и глобальных вычислительных сетей; структуры данных, структуру, состав и свойства информационных процессов,

систем и технологий, базовые и прикладные информационные технологии, классификацию информационных систем, принципы и базовые концепции технологий программирования, основные положения теории баз и хранилищ данных, баз знаний, основные виды и процедуры обработки информации, теорию технологий искусственного интеллекта, алгоритмы машинного обучения и кластерного анализа, состав и структуру инструментальных средств, методологии разработки и проектирования информационных систем.

Должен уметь:

- применять математические методы при решении профессиональных задач повышенной сложности.

Должен иметь навыки:

- построения математической модели профессиональных задач и содержательной интерпретации полученных результатов.

- работать в качестве пользователя с персональным компьютером и программными средствами общего назначения, использовать языки и системы программирования для решения профессиональных задач, разрабатывать информационно-логическую, функциональную и объектно-ориентированную модели информационных систем, модели данных, устанавливать, тестировать и использовать программные компоненты информационных систем, применять информационные технологии при проектировании информационных систем;

- поиска и обмена информацией в локальных и глобальных компьютерных сетях, представления и анализа данных и знаний о предметной области, работы с методами и средствами анализа и реализации проекта информационной системы, владения языками процедурного и объектно-ориентированного программирования, инструментальными средствами обработки информации, интеллектуального анализа данных, владения интеллектуальными технологиями поддержки принятия решений, математического и структурного в виде блок-схем описания алгоритмов обработки информации и машинного обучения.

4. Тематический план

4.1 Математический анализ

- ◆ Теорема Вейерштрасса о максимуме и минимуме непрерывной функции.
- ◆ Формула Тейлора с остаточным членом в форме Лагранжа, с остаточным членом в форме Пеано.
- ◆ Интеграл Римана с переменным верхним пределом. Существование первообразной непрерывной функции. Формула Ньютона-Лейбница.
- ◆ Условие независимости криволинейного интеграла от пути интегрирования. Формула Грина.
- ◆ Функции комплексного переменного: показательная, синус, косинус. Основные свойства.

4.2 Дифференциальные уравнения

1. Уравнения с разделяющимися переменными.
2. Уравнения в полных дифференциалах.
3. Линейные однородные дифференциальные уравнения с постоянными коэффициентами.

4. Свойства решений систем линейных однородных дифференциальных уравнений.
5. Свойства решений систем линейных неоднородных дифференциальных уравнений.
6. Линейные уравнения в частных производных первого порядка.

4.3 Алгебра и геометрия

1. Системы линейных уравнений, равносильные СЛУ, критерий совместности СЛУ, решение СЛУ методом последовательного исключения переменных.
2. Определители, их свойства, разложение определителя по строке или столбцу, правило Крамера.
3. Векторные пространства, их свойства, линейная зависимость и линейная независимость векторов, разложение вектора по базису.
4. Линейные операторы, связь между матрицами линейного оператора относительно различных базисов, собственные векторы и собственные значения, характеристическое уравнение.
5. Прямая на плоскости: способы задания и вывод уравнений.

4.4 Методы вычислений

1. Интерполирование и приближение функций. Интерполяционные многочлены Лагранжа, Ньютона. Сплаины. Построение кубического интерполяционного сплайна.
2. Квадратурные формулы прямоугольников, трапеций, Симпсона, оценки погрешности. Квадратурные формулы Гаусса. Кратные интегралы.
3. Основные задачи линейной алгебры. Метод Гаусса, метод прогонки. Метод простой итерации, метод Гаусса-Зейделя.
4. Приближенные методы решения ОДУ. Метод Эйлера. Модификации метода Эйлера. Методы Рунге-Кутты.
5. Приближенные методы решения уравнений в частных производных.

4.5 Уравнения математической физики

1. Физические задачи, приводящие к уравнениям в частных производных.
2. Классификация уравнений в частных производных второго порядка.
3. Уравнения гиперболического типа. Метод Фурье. Формула Даламбера.
4. Уравнения параболического типа. Функция источника для уравнения теплопроводности.
5. Краевые задачи для уравнения Лапласа.
6. Общие свойства гармонических функций.

4.6 Теория вероятностей и математическая статистика

1. Классификация случайных величин. Числовые характеристики случайных величин.
2. Предельная теорема теории вероятностей.
3. Точечные оценки неизвестных параметров: определения, свойства, методы нахождения.
4. Критерий согласия χ^2 .
5. Формула Байеса.
6. Распределение Гаусса.

4.7 Дискретная математика и математическая логика

1. Булевские функции и способы их представления.
2. Полнота и непротиворечивость логики высказываний.
3. Ограниченные детерминированные функции (ОДФ) и конечные автоматы.

4. Необходимые и достаточные условия однозначности алфавитного кодирования.

5. Рекурсивные и вычислимые (по Тьюрингу) функции.

4.8. Основы программирования

Операционные системы, назначение, выполняемые функции. Принципы управления сетью ЭВМ. Технологические принципы создания программ. Синтаксис и семантика языков программирования. Реализация основных управляющих структур в языках программирования Паскаль и СИ. Парадигмы программирования: процедурная, декларативная, объектно-ориентированная.

4.9. Объектно-ориентированный подход к разработке программ

Простые и сложные программные системы. Иерархия абстракций, концепция ООП. Наследование, сфера действия метода, авторизация доступа, расширяемость объектов. Наследуемые статические методы. Виртуальные методы и полиморфизм. Динамические объекты. Конструкторы и деструкторы.

4.10. Основные понятия структур данных

Три уровня описания структур данных: функциональная спецификация, логическое описание, физическое представление. Концепция типа данных. Создание списка, просмотр списка, включение в список и удаление из списка элементов. Понятие стека. Пакет процедур функциональной спецификации. Бинарные деревья. Сортировки и их классификация.

4.11. Технология программирования

Элементы технологии программирования. Свойства программы: работоспособность, читабельность, эффективность, затраты на разработку, стоимость сопровождения. Жизненный цикл программы. Проекты и способы их представления. Нисходящая и восходящая технология программирования. Метод вертикального слоения, взаимодействие слоев. Определение модуля, программирование модулей. Структурное программирование. Основные управляющие конструкции. Отладка и тестирование. Автономное и комплексное тестирование. Эксплуатационная документация: руководство пользователя, руководство системного программиста.

4.12. Базы и хранилища данных

Понятие базы данных и систему управления базой данных. Основные модели данных: иерархические сетевые, реляционные, объектные и объектно-ориентированные, объектно-реляционные, модель данных SQL. Виды классификаций баз данных. Основы реляционной алгебры. Реляционная модель данных. Принципы нормализации, нормальные формы. Общие сведения о языке SQL. Понятие хранилища данных, принципы организации хранилища.

4.13. Искусственный интеллект

Основные понятия искусственного интеллекта, данные и знания. Направления искусственного интеллекта; экспертные системы, инженерия знаний, нечеткие системы, нейронные сети, генетические алгоритмы, мультиагентные системы. Модели представления знаний: логические модели, основанные на бинарной и нечеткой логике, семантические сети, фреймы. Инструментальные системы обработки знаний, язык Пролог.

4.14. Теория баз данных

Модель «Сущность-связь, понятие, назначение сущности, свойства, связи, степень связи, обязательность/необязательность связи.

Ограничения целостности базы данных на уровне атрибута, кортежа, отношения, базы данных. Ссылочная целостность, потенциальные, первичные, внешние ключи. Ограничение NOT NULL. Реализация ограничений целостности базы данных при помощи триггеров. Декларативные ограничения целостности в языке SQL.

Нормализация отношений базы данных. Функциональная и многозначная зависимости. Нормальные формы отношений БД (1НФ, 2НФ, 3НФ, 4НФ, 5НФ).

Структуры хранения данных. Индексирование данных. Структуры индексов данных. Транзакции. Восстановление данных БД при сбоях. Транзакции и параллелизм. Сериализация транзакций с помощью блокировок. Тупики.

Операции реляционной алгебры. Полнота и замкнутость реляционной алгебры. Реляционная алгебра и язык SQL. Оператор языка SQL SELECT. Соединение отношений (JOIN), агрегатные функции, группировка данных.

4.15 Архитектура ЭВМ и системное ПО

Фон-неймановские принципы реализации ЭВМ.

Формат команд и методы адресации.

Подсистема ввода/вывода. Режимы обмена информацией с внешними устройствами. Иерархия ввода/вывода в вычислительной системе.

Управление памятью. Физическая и виртуальная память. Страничная организация памяти.

Архитектурные способы повышения производительности ЭВМ.

Состояние потоков в многозадачной системе с разделением времени.

Семафоры. Их использование в многопоточных приложениях.

Тупики в многозадачных приложениях.

4.16. Технологии обработки информации и машинное обучение. Общая схема процесса обработки информации. Основные процедуры обработки информации.

Классификация базовых подходов к обработке информации. Байесовская теория решений. Решающее правило на основе минимума условного риска. Решающие правила на основе максимума апостериорной вероятности и функции правдоподобия. Понятие разделяющих функций. Обобщенная структура решающего правила. Оценка ошибок. Распознавание образов, описываемых гауссовскими векторами. Распознавание образов, описываемых произвольными законами распределения. Распознавание образов, описываемых бинарными признаками. Параметрическое обучение в задачах распознавания. Методы максимума правдоподобия и максимума апостериорной вероятности. Подстановочные алгоритмы. Непараметрическое обучение в задачах распознавания. Метод Парзена. Метод К-ближайших соседей. Распознавание образов с помощью функций расстояния (мер близости). Метод опорных векторов. Нелинейные преобразования и спрямляющие пространства. Композиционные алгоритмы, описание и базовые подходы. Композиции на основе баггинга (алгоритм случайный лес). Композиции на основе бустинга (алгоритм adaboost). Кластеризация образов. Критерии эффективности. Алгоритм К-средних. Кластеризация образов. Критерии эффективности. Иерархические процедуры группирования. Основы регрессионного анализа данных.

5. Список рекомендуемой литературы

1. Ильин В. А. Математический анализ : учебник : в 2 ч. / В.А. Ильин, В.А. Садовничий, Б.Х. Сендов ; Моск. гос. ун-т им. М.В. Ломоносова; под ред. А.Н. Тихонова .— М. : Проспект : Изд-во Моск. ун-та, 2006.
2. Тихонов А.Н. Дифференциальные уравнения / А.Н. Тихонов, А.Б. Васильева, А.Г. Свешников ; Моск. гос. ун-т им. М.В.Ломоносова .— Изд. 4-е, стер. — М. : ФИЗМАТЛИТ, 2005 .— 253 с.
3. Бахвалов Н.С. Численные методы / Н. С. Бахвалов, Н. П. Жидков, Г. М. Кобельков ; Моск. гос. ун-т им. М.В. Ломоносова .— 4-е изд. — М. : БИНОМ. Лаб. знаний, 2006 .— 636 с.
4. Ильин В.А. Линейная алгебра и аналитическая геометрия / В.А. Ильин, Г.Д. Ким ; Моск. гос. ун-т им. М.В. Ломоносова .— 3-е изд., перераб. и доп. — М. : Изд-во Моск. ун-та : Проспект, 2007 .— 392 с.
5. Тихонов Андрей Николаевич. Уравнения математической физики / А.Н. Тихонов, А.А. Самарский. — 7-е изд. — М. : Изд-во Моск. ун-та : Наука, 2004 .— 798 с.
6. Гмурман В.Е. Теория вероятностей и математическая статистика / В.Е. Гмурман. — М. : Высш. шк., 2004 .— 478 с.
7. Новиков Ф.А. Дискретная математика для программистов / Ф.А. Новиков .— 3-е изд. — СПб. [и др.] : Питер, 2008 .— 383 с.
8. Шумаков П.В. Фаронов В.В. Delphi 5. Руководство разработчика баз данных – М.: Нолидж, 2000.
9. Райли Д. Абстракция и структуры данных – М.: МИР, 1993.
10. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложений на C++/ Пер. с англ.- М.: Издательство Бином, 1998.
11. Тюкачев Н.А., Свиридов Ю.Т. Delphi 5. Создание мультимедийных приложений. – М.: Нолидж, 2000.
12. Матвеев М.Г. Модели и методы искусственного интеллекта / М.Г. Матвеев, А.С. Свиридов, Н.А. Алейникова – М.: Финансы и статистика, 2008.
13. Сирота, Александр Анатольевич. Методы и алгоритмы анализа данных и их моделирование в MATLAB : [учебное пособие] / А.А. Сирота .— Санкт-Петербург : БХВ-Петербург, 2016 .— 381 с.

6. Образец контрольно-измерительного материала

1. Решить следующую краевую задачу:

$$\begin{cases} u_t - 9u_{xx} = 1 + 2xt + 5t \sin\left(\frac{3\pi x}{2}\right), x \in [0,1], \\ u(0,t) = t, u_x(1,t) = t^2, \\ u(x,0) = 3x. \end{cases}$$

2. Дать математическое описание для алгоритмов кластерного анализа, а именно для алгоритма K-means и алгоритма иерархическое группировки.

3. Неориентированный граф G задан матрицей смежности вершин a: $a[i, j] = 1$, если вершина с номером i – смежная (связана) с вершиной j.

На любом языке программирования реализовать функцию, которая проверяет, является ли граф G связным. Функция (метод) должна в качестве параметра принимать матрицу смежности и возвращать логическое значение (true – граф связный, false – не связный). При необходимости можно реализовать вспомогательные функции (методы).

7. Вариант ответа на КИМ

Решение задания №1.

Чтобы свести все к задаче с однородными граничными условиями, сделаем замену неизвестной функции вида

$$u(x, t) = H(x, t) + \tilde{u}(x, t),$$

где $H(x, t)$ – любая функция, удовлетворяющая неоднородным граничным условиям, т. е.

$$H(0, t) = t, H_x(1, t) = t^2.$$

Будем искать $H(x, t)$ в виде

$$H(x, t) = a(t) + xb(t).$$

Неопределенные коэффициенты найдем с помощью граничных условий:

$$\begin{aligned} H(0, t) &= a(t) = t, \\ H_x(1, t) &= b(t) = t^2. \end{aligned}$$

Таким образом, искомая функция $H(x, t)$ выглядит следующим образом:

$$H(x, t) = t + xt^2.$$

Выполнив замену неизвестной функции

$$u(x, t) = t + xt^2 + \tilde{u}(x, t),$$

получим задачу с однородными граничными условиями

$$\begin{cases} \tilde{u}_t - 9\tilde{u}_{xx} = 5t \sin\left(\frac{3\pi x}{2}\right), x \in [0, 1], \\ \tilde{u}(0, t) = 0, \tilde{u}_x(1, t) = 0, \\ \tilde{u}(x, 0) = 3x. \end{cases}$$

Функцию $\tilde{u}(x, t)$ представим в виде суммы двух компонент

$$\tilde{u}(x,t) = V(x,t) + W(x,t),$$

являющихся решением следующих задач:

$$\begin{cases} V_t - 9V_{xx} = 0, x \in [0,1], \\ V(0,t) = 0, V_x(1,t) = 0, \\ V(x,0) = 3x. \end{cases} \quad \begin{cases} W_t - 9W_{xx} = 5t \sin\left(\frac{3\pi x}{2}\right), x \in [0,1], \\ W(0,t) = 0, W_x(1,t) = 0, \\ W(x,0) = 0. \end{cases}$$

Найдем сначала функцию $V(x,t)$. Для этого поищем решения, представимые в виде произведения $V(x,t) = X(x)T(t)$:

$$XT' - 9X''T = 0.$$

Поделив это равенство на $9XT$, затем разделив переменные, получим

$$\frac{T'}{9T} = \frac{X''}{X} = \text{const.}$$

Константу в правой части для удобства обозначим как $-\lambda^2$. В результате получим следующие два уравнения:

$$X'' + \lambda^2 X = 0,$$

$$T' + 9\lambda^2 T = 0.$$

Решим уравнение для $X(x)$:

$$X(x) = C_1 \cos(\lambda x) + C_2 \sin(\lambda x).$$

С учетом граничных условий $X(0) = 0$, $X'(1) = 0$, получим следующий набор решений:

$$X_n(x) = C_n \sin(\lambda_n x), \lambda_n = \frac{\pi}{2}(1 + 2n), n \in \mathbb{Z}.$$

Среди найденных $X_n(x)$ линейно независимыми будут функции

$$X_n(x) = \sin(\lambda_n x), n = 0, 1, 2, \dots$$

Находим временную компоненту $T(t)$:

$$T_n(t) = A_n e^{-9\lambda_n^2 t}.$$

Умножаем $X_n(x)$ на $T_n(t)$ и складываем:

$$V(x,t) = \sum_{n=0}^{\infty} X_n(x) T_n(t) = \sum_{n=0}^{\infty} A_n e^{-9\lambda_n^2 t} \sin(\lambda_n x).$$

Используем начальные условия $V(x,0) = 3x$:

$$V(x,0) = \sum_{n=0}^{\infty} A_n \sin(\lambda_n x) = 3x.$$

Коэффициенты A_n можно найти, разложив функцию $3x$ в ряд Фурье:

$$A_n = 2 \int_0^1 3x \sin(\lambda_n x) dx = \frac{6 \sin \lambda_n}{\lambda_n^2}.$$

Таким образом, функция $V(x,t)$ выглядит следующим образом:

$$V(x,t) = \sum_{n=0}^{\infty} \frac{6 \sin \lambda_n}{\lambda_n^2} e^{-9\lambda_n^2 t} \sin(\lambda_n x).$$

Теперь найдем функцию $W(x,t)$. Для этого запишем ее в виде ряда Фурье по базису $X_n(x)$:

$$W(x,t) = \sum_{n=0}^{\infty} W_n(t) X_n(x) = \sum_{n=0}^{\infty} W_n(t) \sin(\lambda_n x).$$

Подставим данное выражение в уравнение для $W(x,t)$:

$$\sum_{n=0}^{\infty} W_n'(t) \sin(\lambda_n x) - 9 \sum_{n=0}^{\infty} W_n(t) (\sin(\lambda_n x))'' = 5t \sin\left(\frac{3\pi x}{2}\right).$$

После некоторых упрощений получим следующее равенство:

$$\sum_{n=0}^{\infty} (W_n'(t) + 9\lambda_n^2 W_n(t)) \sin(\lambda_n x) = 5t \sin\left(\frac{3\pi x}{2}\right).$$

Подставим $\lambda_n = \frac{\pi}{2}(1+2n)$ в это соотношение:

$$\sum_{n=0}^{\infty} \left(W_n'(t) + \frac{9\pi^2}{4} (1+2n)^2 W_n(t) \right) \sin\left(\frac{\pi}{2}(1+2n)x\right) = 5t \sin\left(\frac{3\pi x}{2}\right).$$

Приравнивая коэффициенты при одинаковых синусах в левой и правой части, получим следующий набор уравнений:

$$W_n'(t) + \frac{9\pi^2}{4} (1+2n)^2 W_n(t) = 0, n \neq 1,$$

$$W_1'(t) + \frac{81\pi^2}{4} W_1(t) = 5t.$$

С учетом однородных начальных условий $W_n(0) = 0$, при $n \neq 1$ получим только тривиальное решение $W_n(t) = 0$.

Рассмотрим случай $n = 1$. Функцию $W_1(t)$ будем искать как сумму общего решения однородного уравнения $W_1^{oo}(t)$ и частного решения неоднородного $W_1^{qh}(t)$. Функцию $W_1^{oo}(t)$ можно найти с помощью характеристического уравнения. Она выглядит следующим образом:

$$W_1^{oo}(t) = Ce^{-\frac{81\pi^2}{4}t}.$$

Функцию $W_1^{qh}(t)$ можно поискать по виду правой части, т. е. записать как линейную функцию с неопределенными коэффициентами:

$$W_1^{qh}(t) = \alpha + \beta t.$$

Подстановка в исходное уравнение дает следующий результат:

$$W_1^{qh}(t) = \frac{20}{81\pi^2} \left(-\frac{4}{81\pi^2} + t \right).$$

Отсюда находим функцию $W_1(t)$:

$$W_1(t) = Ce^{-\frac{81\pi^2}{4}t} + \frac{20}{81\pi^2} \left(-\frac{4}{81\pi^2} + t \right).$$

С учетом начальных условий $W_1(0) = 0$, последняя формула приобретает следующий вид:

$$W_1(t) = \frac{20}{81\pi^2} \left(\frac{4}{81\pi^2} \left(e^{-\frac{81\pi^2}{4}t} - 1 \right) + t \right).$$

Осталось записать ответ:

$$u(x, t) = t + xt^2 + \sum_{n=0}^{\infty} \frac{6 \sin \lambda_n}{\lambda_n^2} e^{-9\lambda_n^2 t} \sin(\lambda_n x) + \frac{20}{81\pi^2} \left(\frac{4}{81\pi^2} \left(e^{-\frac{81\pi^2}{4}t} - 1 \right) + t \right) \sin\left(\frac{3\pi x}{2}\right),$$

где $\lambda_n = \frac{\pi}{2}(1 + 2n)$.

Решение задания №2.

I. Алгоритм K – means (стандартная реализация) по шагам:

1. Выбираются начальные значения M центров кластеров $m_i, i = \overline{1, M}$, в качестве которых назначаются либо первые образы из исходной выборки X^N (т.е. фактически случайно), либо M наиболее удаленных друг от друга образов выборки (максиминный алгоритм).

Далее в цикле, охватывающем последующие пункты, осуществляются следующие действия.

2. Проводится разбиение $Q_M = \{X^{N_1}, \dots, X^{N_M}\}$ исходной выборки на кластеры, в ходе которого принадлежность каждого образа определяется на основе поиска минимума среди расстояний до установленных ранее центров $m_i, i = \overline{1, M}$. Рассчитывается начальное значение $E_w(X^N, Q_M)$.

3. Проводится перерасчет центров с учетом результатов выполненной в п.2 кластеризации

$$m'_i = \frac{1}{N_i} \sum_{k=1}^{N_i} x^{(k,i)}, i = \overline{1, M}$$

и соответствующий перерасчет функции $E'_w(X^N, Q_M)$.

4. Если разница значений центров на соседних шагах $\delta_{\max} = \max |m_i - m'_i| > \delta_0$, то осуществляется присвоение $m_i = m'_i, i = \overline{1, M}$ и переход к выполнению следующего шага в п. 2. Иначе – останов.

В п.4. δ_0 – порог для фиксации останова итеративного процесса пересчета при условии незначительности изменения центров кластеров на соседних шагах.

II. Агломеративный алгоритм кластеризации (стандартная реализация) по шагам:

1. Устанавливается значение $\tilde{M} = N$ и формируется начальное разбиение на группы (начальная кластеризация) $Q_N^1, Q_N^1 = \{X^1, \dots, X^N\}, X^j = x^{(j)}$.

2. В цикле для выполненной ранее кластеризации Q_{N-k+1}^k и $\tilde{M} = N - k + 1$ с использованием заданной функции расстояния групп находится ближайшая пара групп, имеющая $\min_{i,j} D(X^{N_i}, X^{N_j})$.

3. Проводится слияние выбранных групп $X^{N_i+N_j} = X^{N_i} \cup X^{N_j}$, изъятие X^{N_j} из Q_{N-k+1}^k и образование новой кластеризации Q_{N-k}^k с соответствующими новыми обозначениями групп и присвоением $\tilde{M} = N - k$.

4. Если $\tilde{M} = 1$ (можно $\tilde{M} = M$, если число классов заранее задано) – останов.

Принципиальное значение для подобных алгоритмов имеет выбор функции расстояния для определения близости кластеров. Используются различные варианты, к которым относятся следующие функции:

- расстояние между ближайшими соседями

$$D^{NN}(X^{N_i}, X^{N_j}) = \min_{k,s} d(x^{(k,i)}, x^{(s,j)});$$

- расстояние между самыми удаленными соседями

$$D^{FN}(X^{N_i}, X^{N_j}) = \max_{k,s} d(x^{(k,i)}, x^{(s,j)});$$

- среднее расстояние между образами кластеров

$$D^{BG}(X^{N_i}, X^{N_j}) = \frac{1}{N_i N_j} \sum_{k=1}^{N_i} \sum_{s=1}^{N_j} d(x^{(k,i)}, x^{(s,j)});$$

- расстояние между центрами кластеров $D^{BC}(X^{N_i}, X^{N_j}) = d^2(m_i, m_j)$,

$$m_i = \frac{1}{N_i} \sum_{k=1}^{N_i} x^{(k,i)}, \quad m_j = \frac{1}{N_j} \sum_{s=1}^{N_j} x^{(s,i)};$$

- расстояние Уорда (минимум соответствует максимуму прироста суммы квадратов отклонений до центров кластеров при их объединении)

$$D^{WM}(X^{N_i}, X^{N_j}) = \frac{N_i N_j}{N_i + N_j} d^2(m_i, m_j).$$

Решение задания №3.

// решение на java, вариант обхода в ширину с очередью

```
public static boolean isConnected(int[][] a) {
    int n = a.length;
    boolean[] visited = new boolean[n];
    Queue<Integer> q = new LinkedList<>();
    q.add(0);
    visited[0] = true;
    while (!q.isEmpty()) {
        int i = q.remove();
        for (int j = 0; j < n; j++) {
            if (a[i][j] > 0 && !visited[j]) {
                q.add(j);
                visited[j] = true;
            }
        }
    }
    for (int i = 0; i < n; i++) {
        if (!visited[i]) {
            return false;
        }
    }
    return true;
}
```

```

// решение на c++, вариант рекурсивного обхода в глубину
void visit(vector<vector<int>> &a, vector<int> &v, int i) {
    int n = a.size();
    v[i] = 1;
    for (int j = 0; j < n; j++) {
        if (a[i][j] && !v[j]) {
            visit(a, v, j);
        }
    }
}

bool is_connected(vector<vector<int>> &a) {
    int n = a.size();
    vector<int> v(n); // visited
    visit(a, v, 0);
    return all_of(v.begin(), v.end(), [](int i) { return i; });
}

```

8. Критерии оценки качества поступающего

Экспертная оценка уровня знаний по предлагаемой программе от 0 до 100 баллов. Минимальное количество баллов, подтверждающее успешное прохождение вступительного испытания, составляет 40 баллов.