

МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)

УТВЕРЖДАЮ

Заведующий кафедрой
Математического обеспечения ЭВМ



Г.В.Абрамов
25.04.2022 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Б1.В.14 Алгоритмы биоинформатики

1. Код и наименование направления подготовки/специальности:

02.03.02 Фундаментальная информатика и информационные технологии

2. Профиль подготовки/специализация: Инженерия программного обеспечения

3. Квалификация (степень) выпускника: бакалавр

4. Форма обучения: очная

5. Кафедра, отвечающая за реализацию дисциплины:

Математического обеспечения ЭВМ (МО ЭВМ)

6. Составители программы: Курченкова Т.В., к.т.н., доцент

7. Рекомендована: НМС факультета ПММ, протокол №8 от 15.04.2022 г

(наименование рекомендующей структуры, дата, номер протокола,

отметки о продлении вносятся вручную)

8. Учебный год: 2025/2026

Семестр(ы): 8

9. Цели и задачи учебной дисциплины

Целями освоения учебной дисциплины являются: изучение методов и алгоритмов эффективного вычисления образцов (паттернов) в последовательностях и основ анализа этих алгоритмов, получение навыков проведения отдельных видов исследований в рамках поставленных задач по стандартным методикам, обработки, анализа и оформления результатов исследований и разработок под руководством специалиста более высокой квалификации.

Задачи учебной дисциплины:

- изучение основ современных методов и алгоритмов эффективного вычисления образцов (паттернов) в последовательностях и основы анализа этих алгоритмов, рассмотрение типовых задач обработки больших последовательностей;
- получение навыков проектирования алгоритмов и структур данных, вычисления и использования оценок сложности алгоритмов с целью оптимизации, с учетом требований предметной области и потребностей пользователей;
- проведение эксперимента в соответствии с установленными полномочиями, составление его описание и формулировка выводов, планирование отдельных стадий исследования при наличии общего плана работы;
- получение навыков обработки полученных результатов исследований с использованием стандартных методов (методик);
- применение при обработке данных стандартного и оригинального программного обеспечения;
- представление/оформление результатов лабораторных испытаний в соответствии с действующими технологическими регламентами/требованиями.

10. Место учебной дисциплины в структуре ООП: блок Б1, часть, формируемая участниками образовательных отношений (вариативная).

Студенты должны знать дисциплины: математический анализ, дискретная математика, математическая логика и теория алгоритмов, алгоритмы и анализ сложности, информатика и программирование.

11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями) и индикаторами их достижения:

Код	Название компетенции	Код(ы)	Индикатор(ы)	Планируемые результаты обучения
ПКВ-2	Способен проводить отдельные виды исследований в рамках поставленных задач по стандартным методикам	ПК-2.1	Планирует отдельные стадии исследования при наличии общего плана работы.	знать: критерии оценки алгоритмов; общие концепции теории алгоритмов на строках; методы получения асимптотических оценок. уметь: разрабатывать алгоритмы решения задач обработки больших строк; формулировать задачи оценки сложности; различать классы алгоритмов на строках.
		ПК-2.2	Проводит эксперимент в соответствии с установленными полномочиями, составляет его описание и формулирует выводы	владеть (иметь навык(и)): подходами к исследованию алгоритмов обработки больших строк; математическим аппаратом для оценивания сложности алгоритмов
ПКВ-	Способен	ПКВ-3.1	Обрабатывает	знать: общие концепции теории алгоритмов

3	обрабатывать, анализировать и оформлять результаты исследований и разработок под руководством специалиста более высокой квалификации		полученные результаты исследований с использованием стандартных методов (методик).	на строках; критерии оценки таких алгоритмов; методы получения асимптотических оценок; конкретные алгоритмы.
		ПКВ-3.2	Применяет при обработке данных стандартное и оригинальное программное обеспечение.	уметь: разрабатывать алгоритмы решения задач обработки больших строк; формулировать задачи оценки сложности; различать классы алгоритмов на строках; получать и доказывать оценки сложности; сопоставлять различные алгоритмы на строках, применять при обработке данных стандартное и оригинальное программное обеспечение, оформлять результаты лабораторных испытаний в соответствии с действующими технологическими требованиями.
		ПКВ-3.3	Представляет/оформляет результаты лабораторных испытаний в соответствии с действующими технологическими регламентами/требованиями.	владеть (иметь навык(и)): подходами к исследованию алгоритмов обработки больших строк; математическим аппаратом для оценивания сложности алгоритмов; соответствующим программным обеспечением

12. Объем дисциплины в зачетных единицах/час.(в соответствии с учебным планом) — 2/72.

Форма промежуточной аттестации(зачет/экзамен) зачет.

13. Трудоемкость по видам учебной работы

Вид учебной работы	Трудоемкость	
	Всего	По семестрам
		№ 8
Аудиторные занятия	32	32
в том числе:	лекции	16
	практические	
	лабораторные	16
Самостоятельная работа	40	40
в том числе: курсовая работа (проект)		
Форма промежуточной аттестации (зачет – 0 час. / экзамен – __ час.)		
Итого:	72	72

13.1. Содержание дисциплины

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК *
1. Лекции			
1.1	Введение в алгоритмы биоинформатики. Структура строковых	О последовательностях, алгоритмах биоинформатики и оценках сложности. Основные определения. Определение строковой	https://edu.vsu.ru/course/vie

	последовательностей	последовательности. Грани строк. Z-блоки. Паттерны, классификация паттернов	w.php?id=13152#section-1
1.2	Вычисление частных паттернов Задача о поиске образца	Наивный алгоритм. Алгоритм Кнута–Морриса–Пратта. Алгоритм Бойера–Мура. Алгоритм Карпа–Рабина. Алгоритм Демелки–Бейза–Ятса–Гоннета.	https://edu.vsu.ru/course/view.php?id=13152#section-2
1.3	Вычисление внутренних паттернов. Суффиксное дерево	Определение суффиксного дерева. Наивный алгоритм построения. Алгоритм Мак-Крейта. Алгоритм Укконена. Приложения суффиксных деревьев.	https://edu.vsu.ru/course/view.php?id=13152#section-3
1.4	Внутренние паттерны. Суффиксный массив	Основные определения. Структура суффиксного массива. Приложения суффиксных массивов.	https://edu.vsu.ru/course/view.php?id=13152#section-4
1.5	Частные паттерны. Оптимальное выравнивание строк	Расстояния между строками. Алгоритм Вагнера-Фишера. Оптимальное выравнивание строк.	https://edu.vsu.ru/course/view.php?id=13152#section-5
3. Лабораторные занятия			
3.1	Вычисление частных паттернов Задача о поиске образца	Наивный алгоритм. Алгоритм Кнута–Морриса–Пратта. Алгоритм Бойера–Мура. Алгоритм Карпа–Рабина. Алгоритм Демелки–Бейза–Ятса–Гоннета.	https://edu.vsu.ru/course/view.php?id=13152#section-6
3.2	Вычисление внутренних паттернов. Суффиксное дерево	Определение суффиксного дерева. Наивный алгоритм построения. Алгоритм Мак-Крейта. Алгоритм Укконена. Приложения суффиксных деревьев.	https://edu.vsu.ru/course/view.php?id=13152#section-6
3.4	Частные паттерны. Оптимальное выравнивание строк	Расстояния между строками. Алгоритм Вагнера-Фишера. Оптимальное выравнивание строк.	https://edu.vsu.ru/course/view.php?id=13152#section-6

* заполняется, если отдельные разделы дисциплины изучаются с помощью онлайн-курса. В колонке Примечание необходимо указать название онлайн-курса или ЭУМК. В других случаях в ячейки ставятся прочерки.

13.2. Темы (разделы) дисциплины и виды занятий

№ п/п	Наименование темы (раздела) дисциплины	Виды занятий (количество часов)				Всего
		Лекции	Практические	Лабораторные	Самостоятельная работа	
1	Введение в алгоритмы биоинформатики. Структура строковых последовательностей	2			5	7
2	Вычисление частных паттернов. Задача о поиске образца	4		6	10	18
3	Вычисление внутренних паттернов. Суффиксное дерево	4		6	10	18
4	Внутренние паттерны. Суффиксный массив	2		2	5	11
5	Частные паттерны. Оптимальное выравнивание строк	4		2	10	18
	Итого:	16		16	40	72

14. Методические указания для обучающихся по освоению дисциплины

(рекомендации обучающимся по освоению дисциплины: указание наиболее сложных разделов, работа с конспектами лекций, презентационным материалом, рекомендации по выполнению курсовой работы, по организации самостоятельной работы по дисциплине и др)

- При использовании дистанционных образовательных технологий и электронного обучения выполнять все указания преподавателей по работе на LMS-платформе, своевременно подключаться к online-занятиям, соблюдать рекомендации по организации самостоятельной работы

- Дополнительный материал, рекомендации, а также задания к лабораторным занятиям, к зачету размещаются преподавателем на платформе «Электронный университет» (moodle).
- В течение семестра необходимо выполнить 3 лабораторных задания на компьютере. Работа над лабораторными заданиями должна быть завершена за неделю до окончания курса.

15. Перечень основной и дополнительной литературы, ресурсов интернет, необходимых для освоения дисциплины (список литературы оформляется в соответствии с требованиями ГОСТ и используется общая сквозная нумерация для всех видов источников)

а) основная литература:

№ п/п	Источник
1.	Комлева, Н.В. Структуры и алгоритмы компьютерной обработки данных : учебное пособие / Н.В. Комлева ; Международный консорциум «Электронный университет», Московский государственный университет экономики, статистики и информатики, Евразийский открытый институт. – Москва : Московский государственный университет экономики, статистики и информатики, 2004. – 140 с. : ил., табл., схем. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=93226 (дата обращения: 21.03.2021). – ISBN 5-7764-0400-2. – Текст : электронный.
2.	Мейер, Б. Инструменты, алгоритмы и структуры данных / Б. Мейер. – 2-е изд., испр. – Москва : Национальный Открытый Университет «ИНТУИТ», 2016. – 543 с. : схем., ил. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=429033 (дата обращения: 21.03.2021). – Библиогр. в кн. – Текст : электронный.
3.	Гасфилд Д. Строки, деревья и последовательности в алгоритмах: Информатика и вычислительная биология / Пер. с англ. – СПб: Невский диалект, 2007. – 654 с.
4.	Смит Б. Методы и алгоритмы вычислений на строках / Пер. с англ. – М: Вильямс, 2006. – 496 с.
5.	Томас Х. Алгоритмы: построение и анализ. 3-е издание / Х. Томас, Т. Кормен, Ч. Лейзерсон, Р. Ривест, К.Штайн. – М.: Вильямс, 2013. – 1328с.

б) дополнительная литература:

№ п/п	Источник
6.	Окулов С.М. Алгоритмы обработки строк. – М.: БИНОМ, 2013. – 255 с.
7.	Царёв, Р.Ю. Алгоритмы и структуры данных (CДИО) : учебник / Р.Ю. Царёв, А.В. Прокопенко ; Сибирский федеральный университет. – Красноярск : Сибирский федеральный университет (СФУ), 2016. – 204 с. : ил. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=497016 (дата обращения: 21.03.2021). – Библиогр. в кн. – ISBN 978-5-7638-3388-1. – Текст : электронный
8.	Crochemore, M. and Wojciech, R., <i>Jewels of Stringology</i> , World Scientific, 2002, 310 p

в) информационные электронно-образовательные ресурсы (официальные ресурсы интернет)*:

№ п/п	Ресурс
9.	Курс: Алгоритмы биоинформатики. – Режим доступа: https://edu.vsu.ru/course/view.php?id=13152
10.	Вяххи Н.И. Алгоритмы в биоинформатике. Курс лекций [Электронный ресурс] / Computer Science клуб. – СПб, 2013. – Режим доступа: http://compsciclub.ru/courses/algorithmsbioinformatics .

* Вначале указываются ЭБС, с которыми имеются договора у ВГУ, затем открытые электронно-образовательные ресурсы, онлайн-курсы, ЭУМК

16. Перечень учебно-методического обеспечения для самостоятельной работы (учебно-методические рекомендации, пособия, задачки, методические указания по выполнению практических (контрольных), курсовых работ и др.)

№ п/п	Источник
1.	Курс: Алгоритмы биоинформатики. – Режим доступа: https://edu.vsu.ru/course/view.php?id=13152

17. Образовательные технологии, используемые при реализации учебной дисциплины, включая дистанционные образовательные технологии (ДОТ), электронное обучение (ЭО), смешанное обучение):

При реализации дисциплины могут проводиться различные типы лекций (вводная, обзорная и т.д.), семинарские занятия (проблемные, дискуссионные и т.д.), применяться дистанционные образовательные технологии в части освоения лекционного материала, проведения текущей аттестации, самостоятельной работы по дисциплине или отдельным ее разделам и т.д. При применении ЭО и ДОТ необходимо в п. 15 в) указать используемые ресурсы (см. пример выше)

Дисциплина реализуется с применением электронного обучения и дистанционных образовательных технологий. Для организации занятий рекомендован онлайн-курс «Алгоритмы биоинформатики», размещенный на платформе Электронного университета ВГУ (LMS moodle), а также Интернет-ресурсы, приведенные в п.15в

18. Материально-техническое обеспечение дисциплины:

Лекционная аудитория должна быть оборудована учебной мебелью, компьютером, мультимедийным оборудованием (проектор, экран, средства звуковоспроизведения), допускается переносное оборудование.

Практические занятия должны проводиться в специализированной аудитории, оснащенной учебной мебелью и персональными компьютерами с доступом в сеть Интернет (компьютерные классы, студии), мультимедийным оборудованием (мультимедийный проектор, экран, средства звуковоспроизведения), Число рабочих мест в аудитории должно быть таким, чтобы обеспечивалась индивидуальная работа студента на отдельном персональном компьютере.

Для самостоятельной работы необходимы компьютерные классы, помещения, оснащенные компьютерами с доступом к сети Интернет.

Программное обеспечение:

- ОС Windows
- LibreOffice (свободное и/или бесплатное ПО)
- Microsoft Visual Studio Community Edition (свободное и/или бесплатное ПО)
- Adobe Reader (свободное и/или бесплатное ПО)

19. Оценочные средства для проведения текущей и промежуточной аттестаций

Лабораторные задания и тест в электронном варианте, дистанционные формы общения с преподавателем. Контроль самостоятельной работы реализуется с помощью опросов, тестов, вопросов по темам заданий.

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины:

№ п/п	Наименование раздела дисциплины (модуля)	Компетенция(и)	Индикатор(ы) достижения компетенции	Оценочные средства
1	Введение в алгоритмы биоинформатики. Структура строковых последовательностей	ПКВ-2 ПКВ-3	ПКВ-2.1 ПКВ-3.1	Тест 1 Лабораторные работы
2	Вычисление частных паттернов Задача о поиске образца	ПКВ-2 ПКВ-3	ПКВ-2.2 ПКВ-3.2 ПКВ-3.3	Тест 1 Лабораторные работы
3	Вычисление внутренних паттернов. Суффиксное дерево	ПКВ-2 ПКВ-3	ПКВ-2.1 ПКВ-2.2 ПКВ-3.2	Лабораторные работы
4	Внутренние паттерны. Суффиксный массив	ПКВ-2 ПКВ-3	ПКВ-2.1 ПКВ-2.2	Тест 1

№ п/п	Наименование раздела дисциплины (модуля)	Компетенция(и)	Индикатор(ы) достижения компетенции	Оценочные средства
			ПКВ-3.1 ПКВ-3.2	
5	Частные паттерны. Оптимальное выравнивание строк	ПКВ-2 ПКВ-3	ПКВ-2.1 ПКВ-2.2 ПКВ-3.1 ПКВ-3.2	Тест 1 Лабораторные работы
Промежуточная аттестация форма контроля - зачет				<i>Лабораторные работы, Тест 1 Вопросы к зачету (КИМ)</i>

20 Типовые оценочные средства и методические материалы, определяющие процедуры оценивания

Оценка знаний, умений и навыков, характеризующая этапы формирования компетенций в рамках изучения дисциплины осуществляется в ходе текущей и промежуточной аттестаций.

Текущая аттестация проводится в соответствии с Положением о текущей аттестации обучающихся по программам высшего образования Воронежского государственного университета. Текущая аттестация проводится в форме(ах): устного опроса (индивидуальный опрос); письменных работ (лабораторные работы); тестирования. Критерии оценивания приведены выше.

Промежуточная аттестация проводится в соответствии с Положением о промежуточной аттестации обучающихся по программам высшего образования.

Контрольно-измерительные материалы промежуточной аттестации включают в себя теоретические вопросы, позволяющие оценить уровень полученных знаний, практические задания, позволяющие оценить степень сформированности умений и навыков.

При оценивании используются количественные или качественные шкалы оценок. Критерии оценивания приведены выше.

20.1 Текущий контроль успеваемости

Контроль успеваемости по дисциплине осуществляется с помощью следующих оценочных средств: лабораторные работы.

Лабораторные работы по изучению основ современных методов и алгоритмов эффективного вычисления образцов (паттернов) в последовательностях и основ анализа этих алгоритмов.

Перечень заданий для лабораторных работ:

Реализация наивного алгоритма вычисления всех вхождений паттерна в строку.

Реализация алгоритма Бойера–Мура.

Реализация алгоритма Кнута–Морриса–Пратта.

Реализация алгоритма Карпа–Рабина.

Реализация алгоритма Демелки–Бейза–Ятса–Гоннета

Реализация алгоритма Мак-Крейта.

Реализация алгоритма Укконена.

Реализация алгоритма Вагнера-Фишера.

Описание технологии проведения

Лабораторная работа включает в себя этапы: формализация постановки задачи, выбор структуры данных, описание модели взаимодействия пользователя с системой, кодирование, отладка, тестирование, проведение эксперимента в соответствии с установленными полномочиями, составление его описание и формулировка выводов.

Требования к выполнению заданий (или шкалы и критерии оценивания)

- проведение эксперимента, составление его описания и формулировка выводов, планирование отдельных стадий при наличии общего плана работы;
- обработка полученных результатов исследований с использованием стандартных методик;
- применение при обработке данных стандартного и оригинального программного обеспечения;
- оформление результатов лабораторных заданий в соответствии с действующими технологическими требованиями.

20.2 Промежуточная аттестация

Промежуточная аттестация по дисциплине осуществляется с помощью следующих оценочных средств:

- Тест.
- Собеседование по вопросам к зачету.

Перечень вопросов к зачету:

- Понятие строковой последовательности. Основные определения биоинформатики.
- Грани строк. Z-блоки.
- Паттерны.
- Нормальная форма строки.
- Линдонское слово
- Строковая петля.
- Наивный алгоритм поиска паттерна в строке
- Алгоритм Кнута–Морриса–Пратта.
- Алгоритм Бойера–Мура.
- Алгоритм Карпа–Рабина.
- Алгоритм Демелки–Бейза–Ятса–Гоннета.
- Определение суффиксного дерева. Наивный алгоритм построения.
- Алгоритм Мак-Крейта.
- Алгоритм Укконена.
- Суффиксный массив. Основные определения.
- Поиск в суффиксном массиве
- Расстояния между строками.
- Алгоритм Вагнера-Фишера. Оптимальное выравнивание строк.

Тестовые задания

<https://edu.vsu.ru/course/view.php?id=13152>

Критерии аттестации по итогам освоения дисциплины:

зачтено	Грамотно написаны и сданы задачи лабораторного практикума. Хорошее знание теоретического материала. Умение обосновать высказанное утверждение.
не зачтено	Не сданы задачи лабораторного практикума, незнание теоретического материала.