

Белорусский государственный университет

УТВЕРЖДАЮ

Проректор по учебной работе
и образовательным инновациям

О. Н. Здрок

«30» апреля 2020 г.

Регистрационный № УД- 8800 уч.

АЛГОРИТМЫ ВО ВНЕШНЕЙ ПАМЯТИ

**Учебная программа учреждения высшего образования
по учебной дисциплине для специальности:**

1-31 80 09 Прикладная математика и информатика

Профилизация:

Алгоритмы и системы обработки больших данных

2020 г.

Учебная программа составлена на основе ОСВО 1-31 80 09-2019 и учебного плана G31-072/уч. от 11.04.2019 г.

СОСТАВИТЕЛИ:

П. А. Иржавский — старший преподаватель кафедры дискретной математики и алгоритмики факультета прикладной математики и информатики Белорусского государственного университета;

А. А. Толстик — старший преподаватель кафедры вычислительной математики факультета прикладной математики и информатики Белорусского государственного университета;

К. Ю. Вильчевский — старший преподаватель кафедры дискретной математики и алгоритмики факультета прикладной математики и информатики Белорусского государственного университета.

РЕЦЕНЗЕНТЫ:

Ю. Г. Василевский — кандидат физико-математических наук, начальник отдела разработки мобильных сервисов ООО «ЯндексБел».

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой дискретной математики и алгоритмики
(протокол № 14 от 19 марта 2020 года);

Научно-методическим Советом БГУ
(протокол № 4 от 25 марта 2020 года).

Заведующий кафедрой
дискретной математики и алгоритмики В В. М. Котов

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Цели и задачи учебной дисциплины

Учебная дисциплина «Алгоритмы во внешней памяти» знакомит студентов магистратуры с основными методами разработки и анализа алгоритмов решения задач обработки, хранения и эффективного использования массивов данных и информационных структур, размещенных на твердотельных (SSD) и дисковых (HDD) устройствах.

Цель учебной дисциплины — создание базы для разработки эффективных алгоритмов решения прикладных задач, возникающих при использовании больших объемов информации и формирование у магистрантов умения использовать современные научные достижения в области адаптации эффективных алгоритмов для решения конкретных прикладных.

Задачи учебной дисциплины:

1. Изучение особенностей разработки алгоритмов, использующих различные виды памяти компьютера.
2. Использование приближенных и потоковых алгоритмов при решении математических и прикладных задач.
3. Формирование навыков анализа применимости и адаптации готовых методов и подходов при решении конкретной задачи.

Место учебной дисциплины в системе подготовки специалиста с высшим образованием (магистра).

Учебная дисциплина относится к компоненту учреждения высшего образования и входит **в модуль** «Специализированные алгоритмы и структуры данных», является дисциплиной по выбору.

Программа составлена с учетом **межпредметных связей** с учебными дисциплинами. Основой для изучения учебной дисциплины являются следующие учебные дисциплины первой ступени высшего образования: «Теория алгоритмов», «Теория вероятностей и математическая статистика», «Программирование», «Архитектура компьютеров» и «Модели данных и системы управления базами данных» и дисциплины второй ступени высшего образования «Специальные структуры данных» модуля «Алгоритмические аспекты компьютерных наук», «Вероятностные алгоритмы и структуры данных» модуля «Специализированные алгоритмы и структуры данных». Знания, полученные в учебной дисциплине, используются при изучении дисциплины «Алгоритмы обработки текстов» модуля «Специализированные алгоритмы и структуры данных» и «Информационный поиск» модуля «Приложения машинного обучения».

Требования к компетенциям

Освоение учебной дисциплины «Алгоритмы во внешней памяти» должно обеспечить формирование следующих углубленных профессиональных и специализированных компетенций:

специализированные компетенции:

СК-22. Владеть навыками разработки эффективных алгоритмов, использующих различные виды памяти компьютера;

углубленные профессиональные компетенции:

УПК-4. Оценивать эффективность алгоритмов решения прикладных задач.

В результате изучения дисциплины студент магистратуры должен:

знать:

– современное состояние и тенденции развития основных направлений исследований в области разработки эффективных алгоритмов во внешней памяти;

– особенности разработки эффективных алгоритмов для обработки больших объемов данных;

– специальные способы оценки эффективности алгоритмов обработки и использования данных.

уметь:

– разрабатывать и реализовывать эффективные алгоритмы обработки массивов информации в прикладных задачах;

– оценивать эффективность алгоритмов в ИО-модели;

– применять современные достижения в области информатики и математики для обоснования и верификации математических моделей, методов и алгоритмов обработки больших данных.

владеть:

– приемами построения и анализа математических моделей;

– основными подходами к разработке эффективных алгоритмов обработки и использования данных;

– навыками реализации и использования структур данных (в том числе включенных в программные комплексы) обработки больших объемов информации.

Структура учебной дисциплины

Дисциплина изучается в 3-м семестре. Всего на изучение учебной дисциплины «Алгоритмы во внешней памяти» отведено:

– для очной формы получения высшего образования — 198 часов, в том числе 60 аудиторных часов, из них: лекции — 20 часов (в том числе 10 часов дистанционного обучения), практические занятия — 20 часов (в том числе 10 часов дистанционного обучения), семинарские занятия — 20 часов (в дистанционной форме обучения).

Трудоемкость учебной дисциплины составляет 6 зачетных единицы.

Форма текущей аттестации по учебной дисциплине — экзамен.

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Раздел 1. Алгоритмы обработки данных во внешней памяти

Тема 1.1. Модель вычислений во внешней памяти. Оценка сложности алгоритма

Многоуровневая структура памяти компьютера. Оценка сложности алгоритма, учитывающая число арифметических операций и время доступа к различным видам памяти компьютера.

Тема 1.2. Алгоритмы сортировки во внешней памяти

Особенности сортировки во внешней памяти. Оценка сложности сортировки во внешней памяти. Алгоритм сортировки слиянием и его характеристики.

Тема 1.3. Задача упорядочивания списка

Постановка задачи List-Ranking. Примитив «Слияния упорядоченных списков пар». Детерминированный и недетерминированный подходы к удалению элементов. Восстановление рангов. Использование List-Ranking при обходе деревьев.

Раздел 2. Структуры данных во внешней памяти

Тема 2.1. Деревья во внешней памяти

Деревья поиска. В-деревья поиска и их разновидности. Базовые операции и их трудоемкость. Разреженные В-деревья. Буферизированные В-деревья.

Тема 2.2. Кучи во внешней памяти

Бинарные и d -арные кучи. Алгоритм Дейкстры и поиск в глубину во внешней памяти. Вычисления на ориентированных графах.

Тема 2.3. Алгоритмы обхода графов

Обход в глубину во внешней памяти. Обход в ширину во внешней памяти. Выделение компонент связности. Построение минимального остовного дерева.

Раздел 3. Поточковые алгоритмы

Тема 3.1. Поточковые алгоритмы. Поиск порядковых статистик

Потоки данных. Область применения поточковых алгоритмов. Порядковые статистики. Точные и приближенные алгоритмы поиска порядковых статистик.

Тема 3.2. Приближенный подсчет числа различных элементов

Приближенный алгоритм подсчета числа различных элементов в потоке. Оценка качества приближения. Фильтр Блума для подсчета числа различных элементов в потоке.

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Дневная форма получения образования с применением дистанционных образовательных технологий

№ п/п	Название раздела, темы	Количество аудиторных часов				Количество часов УСР	Форма контроля знаний
		Лекции	Семинарские занятия	Практически е занятия	Иное		
1	Алгоритмы обработки данных во внешней памяти	6	4 (ДО)	6			
1.1	Модель вычислений во внешней памяти. Оценка сложности алгоритма	2		2		Устный опрос	
1.2	Алгоритмы сортировки во внешней памяти	2	2 (ДО)	2		Устный опрос. Доклад	
1.3	Задача упорядочивания списка	2	2 (ДО)	2		Доклад. Отчеты по домашним практическим заданиям с их устной защитой. Контрольная работа № 1	
2	Структуры данных во внешней памяти	4 6 (ДО)	10 (ДО)	4 6 (ДО)			
2.1	Деревья во внешней памяти	2	2 (ДО)	2		Устный опрос. Доклад. Отчеты по домашним практическим заданиям с их устной защитой	
2.2	Кучи во внешней памяти	2 2 (ДО)	4 (ДО)	2 2 (ДО)		Устный опрос. Доклад. Отчеты по домашним практическим заданиям с их устной защитой	
2.3	Алгоритмы обхода графов	4 (ДО)	4 (ДО)	4 (ДО)		Устный опрос. Доклад. Отчеты по домашним практическим заданиям с	

							их устной защитой. Контрольная работа № 2
3	Потоковые алгоритмы	4 (ДО)	6 (ДО)	4 (ДО)			
3.1	Потоковые алгоритмы. Поиск порядковых статистик	2 (ДО)	4 (ДО)	2 (ДО)			Устный опрос. Доклад. Отчеты по домашним практическим заданиям с их устной защитой
3.2	Приближенный подсчет числа различных элементов	2 (ДО)	2 (ДО)	2 (ДО)			Доклад. Отчеты по домашним практическим заданиям с их устной защитой. Коллоквиум

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Перечень основной литературы

1. Veith D., Meyer U. C., Schnitg G.. Graph Algorithms for Approximate and Dynamic Settings in the External-memory Model. Johann Wolfgang Goethe-Universität Frankfurt am Main, 2019
2. Кнут, Д. Э.. Искусство программирования, том 3. Сортировка и поиск, 2-е изд.: Пер. с англ. — М.: ООО «И.Д. Вильямс», 2007. — 832 с.
3. Соболев, С. А. Сборник задач по теории алгоритмов. Структуры данных: учеб.-метод. пособие / С.А. Соболев [и др.] — Минск : БГУ, 2020. — 159 с.

Перечень дополнительной литературы

4. Кормен Т. Х., Лейзерсон Ч. И., Ривест Р. Л., Штайн К. Алгоритмы: построение и анализ. — 2-е издание: Пер. с англ. — М.: Вильямс, 2005. — 1296 с.
5. В. М. Котов, Е. П. Соболевская, А. А. Толстиков. «Алгоритмы и структуры данных»: учеб. пособие. Минск: БГУ — 2011. — 267 с. — (Классическое университетское издание).

Перечень рекомендуемых средств диагностики и методика формирования итоговой оценки

Для диагностики компетенции в рамках учебной дисциплины рекомендуется использовать следующие формы:

1. Устная форма: устный опрос, коллоквиум, выступление с докладом на семинаре.
2. Письменная форма: контрольные работы.
3. Устно-письменная форма: отчеты по домашним практическим заданиям с их устной защитой, оценивание на основе проектного метода.

Формой текущей аттестации по дисциплине «Алгоритмы во внешней памяти» учебным планом предусмотрен — экзамен.

При формировании итоговой оценки используется рейтинговая оценка знаний студента, дающая возможность проследить и оценить динамику процесса достижения целей обучения. Рейтинговая оценка предусматривает использование весовых коэффициентов для текущего контроля знаний студентов по дисциплине.

Примерные весовые коэффициенты, определяющие вклад текущего контроля знаний в рейтинговую оценку (формирование оценки за текущую успеваемость):

- отчёты по практическим домашним заданиям с их устной защитой — 40 %;
- контрольные работы — 20 %;
- коллоквиум — 10 %;
- устный опрос — 10%;
- выступление с докладом на семинаре — 20%.

Рейтинговая оценка по дисциплине рассчитывается на основе оценки текущей успеваемости и экзаменационной оценки с учетом их весовых коэффициентов Вес оценка по текущей успеваемости составляет 30 %, экзаменационная оценка — 70 %.

Примерная тематика семинарских занятий

Занятие № 1. Модель вычислений во внешней памяти. Оценка сложности алгоритма. Алгоритмы сортировки во внешней памяти

Занятие № 2. Задача упорядочивания списка.

Занятие № 3. Деревья во внешней памяти.

Занятие № 4. Структура данных COLA.

Занятие № 5. Неосведомлённые о кэш-памяти алгоритмы. Задача о транспонировании матрицы.

Занятие № 6. Неосведомлённые о кэш-памяти алгоритмы. Задача о перемножении матриц.

Занятие № 7. Неосведомлённые о кэш-памяти алгоритмы. Поиск ближайшего элемента в неизменном множестве.

Занятие № 8. Поточковые алгоритмы.

Занятие № 9. Поиск порядковых статистик. Поиск медианы за один проход. Поиск порядковой статистики методом Манро – Паттерсона.

Занятие № 10. Приближенные алгоритмы решения прикладных задач и способы их оценки. Поиск часто встречающихся элементов

Описание инновационных подходов и методов к преподаванию учебной дисциплины

При организации образовательного процесса большинства практических занятий используется *практико-ориентированный подход*, который предполагает освоение содержания учебного материала через решение практических задач, а также приобретение навыков эффективного выполнения разных видов профессиональной деятельности.

Кроме этого, при организации образовательного процесса используется комбинация *методов группового обучения, проектного обучения и учебной дискуссии*. Комбинация методов предполагает: ориентацию на создание

идей, приобретение навыков для решения исследовательских, творческих и коммуникационных задач, появление нового уровня понимания изучаемой темы, применение знаний (теорий, концепций) при решении проблем, определение способов их решения.

Методические рекомендации по организации самостоятельной работы обучающихся, подготовка к экзамену

Для организации самостоятельной работы студентов магистратуры по учебной дисциплине следует использовать информационно-коммуникационные технологии:

образовательный портал БГУ <https://edufpmi.bsu.by>;

образовательный портал InsightRunner <https://acm.bsu.by>;

систему AnyTask <https://anytask.org/school/bsu>;

разместить в сетевом доступе комплекс учебных и учебно-методических материалов (учебно-программные материалы, учебное издание для теоретического изучения дисциплины, презентации лекций, методические указания к практическим занятиям, электронные версии домашних заданий, материалы текущего контроля и текущей аттестации, позволяющие определить соответствие учебной деятельности обучающихся требованиям образовательных стандартов высшего образования и учебно-программной документации, в том числе вопросы для подготовки к экзамену, задания, вопросы для самоконтроля, список рекомендуемой литературы, информационных ресурсов и др.).

Примерный перечень вопросов к экзамену

1. Модель вычислений во внешней памяти.
2. Физика жёсткого диска. Теоретически оптимальный размер блока.
3. Сортировка слиянием во внешней памяти.
4. Сортировка распределением во внешней памяти.
5. Поразрядная сортировка во внешней памяти.
6. Ранжирование списка во внешней памяти.
7. Поиск большого независимого множества в графе-цепи во внешней памяти.
8. Эйлеров обход дерева во внешней памяти.
9. Динамическое программирование по дереву во внешней памяти.
10. (A, B) -дерево.
11. Буферизованное дерево.
12. Куча во внешней памяти.
13. Динамическое программирование с большим числом состояний (обработка наперёд).
14. Структура данных CO_LA, добавление элемента.
15. Структура данных CO_LA, поиск элементов из интервала.

16. Алгоритм построения остова в графе во внешней памяти.
17. Алгоритм построения остова минимального веса в графе во внешней памяти.
18. Алгоритм поиска в ширину в неориентированном графе во внешней памяти.
19. Стек и очередь во внешней памяти.
20. Понятие кэш-памяти. Стратегии вытеснения данных. Структура кэш-памяти процессора.
21. Неосведомлённые о кэш-памяти алгоритмы. Задача о транспонировании матрицы.
22. Неосведомлённые о кэш-памяти алгоритмы. Задача о перемножении матриц.
23. Неосведомлённые о кэш-памяти алгоритмы. Поиск ближайшего элемента в неизменном множестве.
24. Фильтр Блума.
25. Поточные алгоритмы. Моделирование случайного подмножества малой мощности.
26. Поточные алгоритмы. Поиск порядковой статистики методом Манро – Паттерсона.
27. Поточные алгоритмы. Поиск медианы за один проход.
28. Поточные алгоритмы. Поиск часто встречающихся элементов.
29. Поточные алгоритмы. Оценка на кратность часто встречающихся элементов.
30. Поточные алгоритмы. Оценка числа различных элементов.

Рекомендуемая тематика контрольных работ

- 1) Контрольная работа № 1. *Алгоритмы обработки данных во внешней памяти.*
- 2) Контрольная работа № 2. *Структуры данных во внешней памяти.*

Текущий контроль знаний проводится в соответствии с учебно-методической картой дисциплины.

ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ УВО

Название учебной дисциплины, с которой требуется согласование	Название Кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
Алгоритмы обработки текстов	Дискретной математики и алгоритмики	Нет	Оставить содержание учебной дисциплины без изменения, (протокол № 14 от 19 марта 2020 г.)
Информационный поиск	Дискретной математики и алгоритмики	Нет	Оставить содержание учебной дисциплины без изменения, (протокол № 14 от 19 марта 2020 г.)

ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ

на ____ / ____ учебный год

№№ Пп	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры дискретной математики и алгоритмики (протокол № ____ от _____ 20__ г.)

Заведующий кафедрой

(ученая степень, звание)

(подпись)

(И. О. Фамилия)

УТВЕРЖДАЮ

Декан факультета

(ученая степень, звание)

(подпись)

(И. О. Фамилия)