

Задание 2

Тема: Хеширование и организация быстрого поиска данных.

Цель: Получить навыки по разработке хеш-таблиц и их применении.

Задание

1. Разработайте приложение, которое использует хеш-таблицу для организации прямого доступа к записям файла, структура записи которого приведена в варианте.

Разработайте и реализуйте функции для операций:

- 1) Хеш-функцию (метод определите сами).
 - 2) Прочитать запись из файла и вставки запись в таблицу (запись включает: ключ и номер записи с этим ключом в файле).
 - 3) Удалить запись из таблицы и соответственно из файла.
 - 4) Найти запись с заданным ключом в файле, используя хеш-таблицу.
 - 5) Выполнить рехеширование.
2. Разработайте такие тесты, чтобы возникли коллизии.
3. Разработайте такие тесты, чтобы требовалось рехеширование.
4. Заполните файл большим количеством записей. Определите время чтения записи с заданным ключом для первой записи файла, для последней и где-то в середине. Убедитесь (или нет), что время доступа для всех записей одинаково.
5. Выведите список индексов, которые формируются при вставке элементов в таблицу.

Рекомендация. Файл используйте двоичный.

Выбор номера варианта: (остаток от деления номера студента в списке журнала % 17)+1, .

Составить отчет, включив в него

- 1) Ответы на вопросы
 1. Расскажите о назначении хеш-функции.
 2. Что такое коллизия?
 3. Что такое «открытый адрес» по отношению к хеш-таблице?
 4. Как в хеш-таблице с открытым адресом реализуется коллизия?
 5. Какая проблема, может возникнуть после удаления элемента из хеш-таблицы с открытым адресом и как ее устранить?
 6. Что определяет коэффициент нагрузки в хеш-таблице?
 7. Что такое «первичный кластер» в таблице с открытым адресом?
 8. Как реализуется двойное хеширование?
 - 2) Выполните разработку программы в соответствии с технологией разработки и включите в отчет
- Образец отчета по разработке в приложении 1

ВАРИАНТЫ

№	Тип хеш-таблицы (способ реализации коллизий)	Структура записи файла(ключ – подчеркнутое поле)
1	С открытой адресацией (увеличение на 1)	Читательский абонемент: <u>номер читательского</u> - целое пятизначное число, ФИО, Адрес
2	С открытой адресацией (увеличение на номер выполняемого подбора)	Счет в банке: <u>номер счета</u> 7 разрядное число, ФИО, Адрес
3	С открытой адресацией (двойное хеширование)	Владелец телефона: <u>номер телефона</u> – последовательность символов, адрес
4	Цепное хеширование	Владельцев автомобилей. <u>номер машины</u> , марка, сведения о владельце.
5	Цепное хеширование	Пациент поликлиники: <u>номер карточки</u> , код хронического заболевания, Фамилия лечащего врача
6	Цепное хеширование	Товар: название, <u>код</u> – шестизначное число
7	Цепное хеширование	Специализация вуза: <u>код специальности</u> , название вуза
8	Открытый адрес (двойное хеширование)	Книга: <u>ISBN</u> – двенадцатизначное число, Автор, Название
9	Цепное хеширование	Страховой полис: <u>номер</u> , компания, фамилия владельца
10	Открытый адрес (увеличение на 1)	Англо – русский словарь: <u>английское слово</u> , русское слово
11	Открытый адрес (двойное хеширование)	Железнодорожная справка: <u>номер поезда</u> , <u>пункт отправления</u> , пункт назначения, время отправления
12	Цепное хеширование	Регистрация малого предприятия: <u>номер лицензии</u> , название, учредитель
13	Открытый адрес(двойное хеширование)	Студент: <u>номер зачетной книжки</u> , номер группы ФИО
14	Цепное хеширование	Справочная межгорода: <u>код города</u> , название города
15	Открытый адрес (увеличение на 1)	Найти и поздравить друга: <u>дата рождения</u> , имя
16	Цепное хеширование	Частотный словарь: слово, количество вхождений в текст
17	Открытый адрес	Учет товара: <u>артикул</u> , название товара, количество в

	(увеличение на 1)	наличии (штук).
--	-------------------	-----------------

Приложение 1 Образец оформления отчета по разработке приложения

Условие задачи

Разработать приложение, которое использует хеш-таблицу для организации прямого доступа к записям файла, структура записи которого приведена в варианте.

Задание варианта

№	Метод хеширования (способ реализации коллизий)	Структура записи файла (ключ – подчеркнутое поле)
	Цепное хеширование	Частотный словарь: <u>слово</u> , количество вхождений в текст

1. Постановка задачи

Разработайте приложение, которое использует хеш таблицу для организации прямого доступа к записям файла, структура записи которого приведена в варианте.

Дано.

Хеш-таблица для реализации коллизий по методу *Цепное хеширование*.

Файл двоичный с записями фиксированной длины.

Структура записи файла

слово, количество вхождений слова в текст

Результат.

Приложение, выполняющее операции

Управление хеш-таблицей: вставить ключ в таблицу, удалить ключ из таблицы, найти ключ в таблице, рехешировать таблицу

Управление файлом:

1) посредством хеш-таблицы: считать запись из файла при поиске записи

2) добавить запись в файл, удалить запись из файла, прочитав запись файла по заданному номеру записи.

2. Подход к решению

1) Хеш-таблица — класс

2) Структура элемента цепной хеш-таблицы: *указатель на вершину списка (цепочки)*

3) Структура элемента линейного однонаправленного списка: *ключ, ссылка на запись в файле (это адрес места размещения записи/порядковый номер записи в файле), ссылка на следующий элемент списка. Новый элемент вставляется в начало списка.*

4) Методы хеш-таблицы: вставить ключ в таблицу, удалить ключ из таблицы, найти ключ в таблице, рехешировать таблицу, вывод хеш-таблицы в консоль.

5) Файл двоичный из записей фиксированного размера.

Структура записи файла:

слово — текст размером 50 символов,

количество вхождений — целое положительное число

признак логического удаления – логического типа

Операции по управлению файлом:

- добавить запись в файл: занести запись в файл, вставить сведения о записи в хеш-таблицу

- удалить запись из файла: удалить из хеш-таблицы и из файла

- прочитать запись из указанной позиции (номеру/смещению)

3. Алгоритмы операций на псевдокоде:

- вставка в таблицу

- поиск записи по ключу в таблице и возвращение ссылки на запись в файле

- удаление элемента из хеш-таблицы

- вставка записи по ключу

4. Код приложения

5. Скриншот результатов выполнения операций с хеш-таблицей.

6. Скриншоты содержания файла и хеш-таблицы