Министерство образования Республики Беларусь

Учреждение образования

“Белорусский государственный университет

информатики и радиоэлектроники”

Факультет информационных технологий и управления

Кафедра интеллектуальных информационных технологий

**Отчет**

**по лабораторной работе №6**

**по курсу**

**«Аппаратное обеспечение интеллектуальных систем»**

**На тему: Моделирование хеш-таблиц**

Выполнил Носко А.Ю.

студент группы 921702

Проверил Захаров В.В.

Доцент кафедры ИИТ

Минск 2020

**Цель:**

Освоение навыков построения и проверки таблиц хеширования.

**Задание:**

Разработать и проверить программу, обеспечивающую формирование хеш-таблицы по ключевым словам, и выполнение различных операций с этой таблицей – включение в таблицу новых строк, поиск информации в таблице по ключевым словам, удаление строк из таблицы.

**Вариант 20**

Предметная область записываемых ключевых слов - Армия, вооружение.

**Ход и результаты работы:**

1. Определяем перечень ключевых слов. Каждому ключевому слову в таблице, являющимся названием терминов из темы «Армия, Вооружение», ставится в соответствие определение данного термина.
2. Создаем таблицу на 20 строк. Структура ячейки (строки) хеш-таблицы:
3. ID – ключевое слово;
4. V – числовое значение ключевого слова; Применяем способ определения значения ключевого слова (V) по его первым двум буквам. Используем русский алфавит, буквам присваиваем соответствующие номера.
5. h(V) – хеш-адрес, вычисленный по числовому значению V ключевого слова; Применяем метод деления. Хеш-функция имеет вид:

h(V)=V modH, где

* H - общее количество строк (ячеек) в таблице;
* V modH – остаток от деления V на H;

1. C – флажок коллизий; Первоначально записан «0». После первой коллизии в данной ячейке флажок переводится в «1».
2. U – флажок «занято»; «1»– если в ячейке хранится информация, «0» – если ячейка свободна или хранившийся в ней элемент помечен как вычеркнутый из таблицы;
3. T – терминальный флажок. Отражает конец последовательности пробинга;
4. D – флажок вычеркивания; Первоначально записан «0». «1» – строка удаляется.
5. L – флажок связи; Значение «0» этого флажка означает, что искомые данные хранятся в той же ячейке таблицы, если же флажок связи равен «1», то соответствующее информационное поле нужно рассматривать как хеш-связь.
6. Po - указатель следующей записи в цепочке коллизий, если таковая имеется, иначе Pо = ID;
7. Pi – данные;
8. Заносим в таблицу 11 понятий из предметной области. В дальнейшем с хеш-таблицей можно выполнять различные операции: поиск информации по ключевому слову, запись в таблицу новых строк, удаление отдельных строк из таблицы.
9. Когда на одну ячейку претендуют два или более ключа возникает коллизия. Выбираем способ обработки коллизий - размещение резервных ячеек в хеш-таблице (внутренняя адресация).
10. Способ пробинга при выборе резервных ячеек при использовании внутренней адресации – линейный. К хеш-адресу последовательно прибавляется по единице до тех пор, пока не обнаружится пустая ячейка.
11. Вид первоначальной таблицы, содержащей 11 элементов:



1. Вид таблицы после удаления ячейки с ключевым словом Атака:



1. Вывод по запросу поиска ключевого слова Генерал:

C:\Users\usn55\Desktop\АОИС_лабы\LR6\3.PNG

**Выводы:**

В ассоциативных способах поиска информации доступ к любой информационной единице (записи, документу, массиву или элементу данных и т.д.) осуществляется:

– либо по специальному ключу;

– либо по некоторому фрагменту самой информационной единицы.

Поиск может быть реализован как аппаратно, так и программно. Программные методы называют методами хеширования.

Функция, которая трансформирует ключ в некоторый индекс, называется хеш-функцией.

Все допустимые ключевые слова образуют так называемое пространство имен, а все адреса памяти, в которые преобразуются ключевые слова, – адресное пространство.

Получаемые из ключевых слов методом хеширования адреса называются хеш-адресами.

Хеш-таблица, или таблица хеширования (ТХ) – это таблица с адресацией, задаваемой хеш-функцией.

Формирование хеш-функции проводится в два этапа:

1. выбор способа перевода ключевых слов в числовую форму;
2. выбор алгоритма преобразования числовых значений в хеш-адреса.

В данной лабораторной работе применяем способ определения значения ключевого слова (V) по его первым двум буквам. Используем русский алфавит, буквам присваиваем соответствующие номера.

Для преобразование числовых значений в хеш-адреса применяются различные методы хеширования: метод деления, метод умножения, метод квадрата, метод извлечения битов, переход к новому основанию. В данной лабораторной работе применяем метод деления. Данный метод является наиболее распространенным и обеспечивает достаточно равномерное распределение хеш-адресов в пределах любой заданной области памяти

Ситуация, когда два или более ключа ассоциируются с одной и той же позицией, называется коллизией (конфликтом) при хешировании. В этом случае для размещения второй и последующих записей необходимо использовать резервные ячейки памяти, которые размещаются либо в самой таблице хеширования (внутренняя адресация), либо в специальной области памяти (области переполнения). В данной лабораторной работе выбираем способ обработки коллизий - размещение резервных ячеек в хеш-таблице (внутренняя адресация).

Для поиска свободных резервных ячеек при внутренней адресации применяется процедура пробинга (линейного, квадратичного или случайного). В данной лабораторной работе выбираем способ пробинга при выборе резервных ячеек – линейный. Алгоритм линейного типа хорошо работает в начале заполнения таблицы, однако по мере заполнения процесс замедляется. Однако он достаточно прост и эффективен.