Lab 2

Michael Morikawa & James Daza

March 5, 2020

Lab Questions

Question 1 What are some effective techniques to generate an integer value for a key of data type string?
Polynomial hashing and cyclic shift are the most effective.

Question 2 Whate are some examples of compression functions?
Division $f(x)\%N$ **MAD:** $(a \cdot f(x) + b)\%N)$

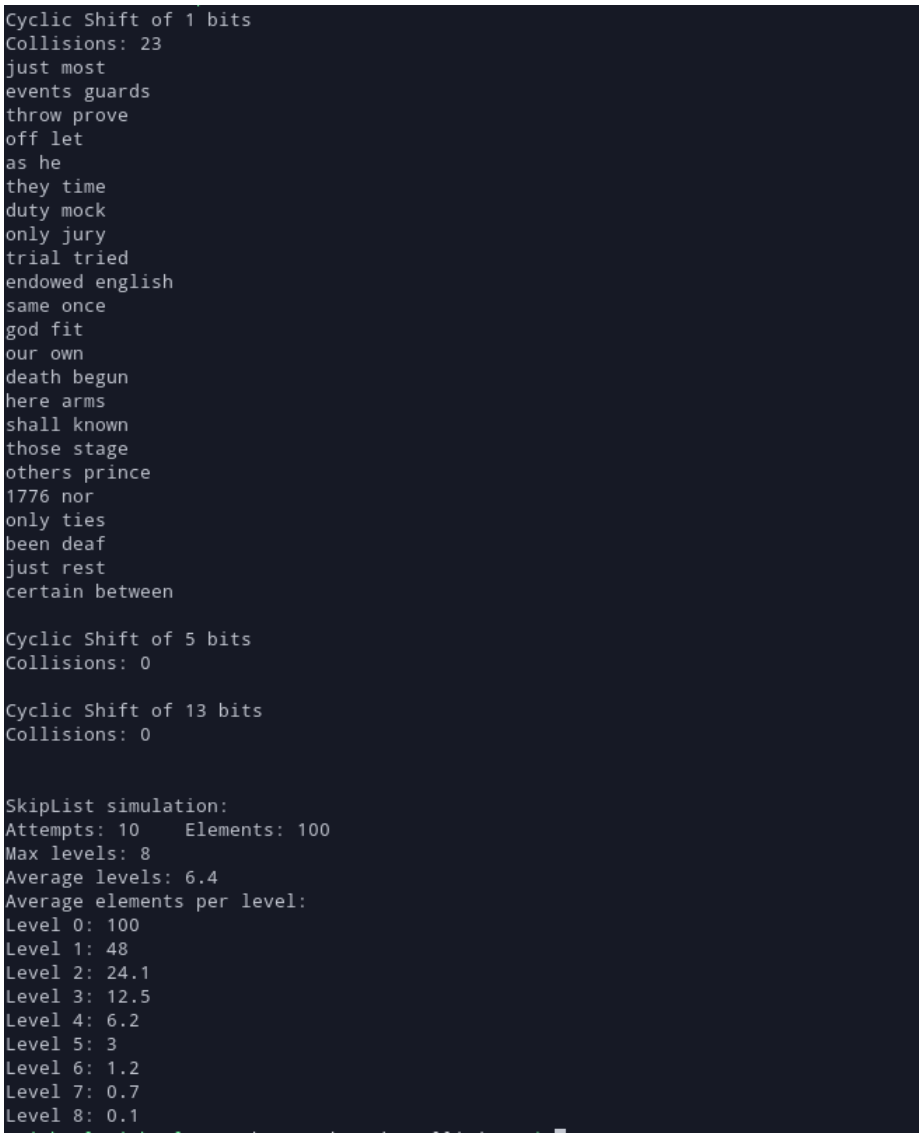
Source Code

```
#include <algorithm>
#include <fstream>
#include <iostream>
#include <map>
#include <string>
#include <time.h>
#include <vector>

void allLowerCase(std::string& word)
{
    for (unsigned int i = 0; i < word.size(); i++)
    {
        word[i] = std::tolower(word[i]);
    }
}

int hashCode(std::string word, int shift)
{
    unsigned int x = 0;
    for (unsigned int i = 0; i < word.size(); i++)
    {
        x = (x << shift) | (x >> 32 - shift);
        x += (unsigned int)word[i];
    }
    return (int)x;
}

void collisionTest(std::vector<std::string> wordList, int shift)
{
    std::map<std::string, int> encounteredWords;
    std::map<int, std::string> collisionTest;
    int code = 0;
    int totalCollisions = 0;
    std::vector<std::string> collidedWords;
    for (std::string word : wordList)
        if (encounteredWords.find(word) == encounteredWords.end())
```

```
        {
            encounteredWords.insert(std::pair<std::string, int>(word, word.size()));
            code = hashCode(word, shift);
            if (collisionTest.find(code) == collisionTest.end())
            {
                collisionTest.insert(std::pair<int, std::string>(code, word));
            }
            else
            {
                totalCollisions++;
                collidedWords.push_back(collisionTest.find(code)->second);
                collidedWords.push_back(word);
            }
        }
    std::cout << "Cyclic Shift of " << shift << " bits\n"
              << "Collisions: " << totalCollisions << '\n';
    for (unsigned int i = 0; i < collidedWords.size(); i += 2)
    {
        std::cout << collidedWords[i] << ' ' << collidedWords[i + 1] << '\n';
    }
}

int main()
{
    std::ifstream infile;
    infile.open("usdeclarPC.txt");
    std::vector<std::string> wordList;
    std::string word;
    //collision testing
    while (infile >> word)
    {
        //Removes anything that is not a number or an letter
        word.erase(std::remove_if(word.begin(), word.end(), ispunct), word.end());
        allLowerCase(word);
        wordList.push_back(word);
    }
    infile.close();
    collisionTest(wordList, 1);
    std::cout << '\n';
    collisionTest(wordList, 5);
    std::cout << '\n';
    collisionTest(wordList, 13);
    std::cout << '\n';
    //Skiplist simulation

    int elements = 100;
    int attempts = 10;
    int sumLevel = 0;
    int levels[100] = {0};
    int currentLevel = 0;
    int topLevel = 0;
    int maxLevel = 0;
    std::srand(time(NULL));
    for (int i = 0; i < attempts; i++)
    {
        for (int i = 0; i < elements; i++)
        {
```

```
            levels[0]++;
            while (rand() % 2)
            {
                currentLevel++;
                levels[currentLevel]++;
            }
            topLevel = std::max(topLevel, currentLevel);
            currentLevel = 0;
        }
        sumLevel += topLevel;
        maxLevel = std::max(topLevel, maxLevel);
        topLevel = 0;
    }
    std::cout << "\nSkipList simulation:\n"
              << "Attempts: " << attempts << "    Elements: " << elements << "\n"
              << "Max levels: " << maxLevel
              << "\nAverage levels: " << (float)sumLevel / attempts
              << "\nAverage elements per level:\n";

    for (int i = 0; i < 100; i++)
    {
        if (levels[i] == 0)
        {
            break;
        }
        else
        {
            std::cout << "Level " << i << ": " << (float)levels[i] / attempts << '\n';
        }
    }
}
```

Output


Cyclic Shift of 1 bits
Collisions: 23
just most
events guards
throw prove
off let
as he
they time
duty mock
only jury
trial tried
endowed english
same once
god fit
our own
death begun
here arms
shall known
those stage
others prince
1776 nor
only ties
been deaf
just rest
certain between
Cyclic Shift of 5 bits
Collisions: 0
Cyclic Shift of 13 bits
Collisions: 0
SkipList simulation:
Attempts: 10 Elements: 100
Max levels: 8
Average levels: 6.4
Average elements per level:
Level 0: 100
Level 1: 48
Level 2: 24.1
Level 3: 12.5
Level 4: 6.2
Level 5: 3
Level 6: 1.2
Level 7: 0.7
Level 8: 0.1