

Script started on 2024-12-05 12:37:02-06:00 [TERM="xterm-256color" TTY="/dev/pts/6"]
e_prevost@ares:~/Project_2\$ pwd
/home/students/e_prevost/Project_2
e_prevost@ares:~/Project_2\$ cat concordance.info

Robert Prevost

CSC 122 W01

concordance project

A text analysis function that analyzes word usage in text files.

Base Level: Level 6

Total Level: Level 6

*****e_prevost@ares:~/Project_2\$ cat data.txt
When, in the course of human events, it becomes necessary for one people to dissolv

We hold these truths to be self-evident, that all men are created equal, that they

He has refused his assent to laws, the most wholesome and necessary for the public

He has forbidden his governors to pass laws of immediate and pressing importance, i

He has refused to pass other laws for the accommodation of large districts of peop

He has called together legislative bodies at places unusual, uncomfortable, and di

He has dissolved representative houses repeatedly, for opposing with manly firmnes

He has refused for a long time, after such dissolutions, to cause others to be ele

He has endeavored to prevent the population of these States; for that purpose obst

He has obstructed the administration of justice, by refusing his assent to laws fo

He has made judges dependent on his will alone, for the tenure of their offices, ar

He has erected a multitude of new offices, and sent hither swarms of officers to h

He has kept among us, in times of peace, standing armies, without the consent of o

He has affected to render the military independent of and superior to the civil po

He has combined with others to subject us to a jurisdiction foreign to our constit

For quartering large bodies of armed troops among us:

For protecting them, by a mock trial, from punishment for any murders which they sl

For cutting off our trade with all parts of the world:

For imposing taxes on us without our consent:

For depriving us, in many cases, of the benefits of trial by jury:

For transporting us beyond seas to be tried for pretended offences:

For abolishing the free system of English laws in a neighboring province, establishi

For taking away our charters, abolishing our most valuable laws, and altering fundi

For suspending our own legislatures, and declaring themselves invested with power t

He has abdicated government here, by declaring us out of his protection, and waging

He has plundered our seas, ravaged our coasts, burnt our towns, and destroyed the

He is, at this time, transporting large armies of foreign mercenaries to complete t

He has constrained our fellow-citizens, taken captive on the high seas, to bear arr

He has excited domestic insurrections amongst us, and has endeavored to bring on tl

In every stage of these oppressions we have petitioned for redress in the most humil

Nor have we been wanting in attentions to our British brethren. We have warned ther

We, therefore, the Representatives of the United States of America, in General Cong

concordance.cpp:

```
1 #include <iostream>
2 #include <fstream>
3 #include <string>
4 #include <map>
5 #include <vector>
6 #include <set>
7 #include <sstream>
```

```
8 #include <algorithm>
9 #include <cctype>
10 #include <climits>
11 using namespace std;
12
13 struct Location {
14     std::size_t line;
15     std::size_t position;
16 };
17
18 class Concordance {
19 private:
20     map<string, vector<Location>> index;
21     set<string> trivialWords;
22     string currentFile;
23
24     string cleanWord(const string& word) {
25         string cleaned;
26         for (unsigned char c : word) {
27             if (isalnum(c)) {
28                 cleaned += static_cast<char>(tolower(static_cast<unsigned c
29             }
30         }
31         return cleaned;
32     }
33
34 public:
35     Concordance() : index(), trivialWords{"a", "an", "the", "in", "on", "a
36
37     void buildIndex(const string& filename) {
38         currentFile = filename;
39         index.clear();
40         ifstream file(filename);
41         string line;
42         std::size_t lineNum = 0;
43
44         while (getline(file, line)) {
45             lineNum++;
46             istringstream iss(line);
47             string word;
48             std::size_t position = 0;
49
50             while (iss >> word) {
51                 string cleanedWord = cleanWord(word);
52                 if (!cleanedWord.empty() && trivialWords.find(cleanedWord)
53                     Location loc = {lineNum, position};
54                     index[cleanedWord].push_back(loc);
55                 }
56                 position++;
57             }
58         }
59     }
60
61     vector<pair<size_t, string>> searchWord(const string& searchWord) {
```

```

62     vector<pair<size_t, string>> results;
63     string word = cleanWord(searchWord);
64
65     if (index.find(word) == index.end() || !hasIndex()) {
66         return results;
67     }
68
69     ifstream file(currentFile);
70     if (!file.is_open()) {
71         return results;
72     }
73
74     vector<string> lines;
75     string line;
76     while (getline(file, line)) {
77         lines.push_back(line);
78     }
79
80     for (const Location& loc : index[word]) {
81         if (loc.line > 0 && loc.line <= lines.size()) {
82             string context = lines[loc.line - 1];
83             istringstream iss(context);
84             string temp;
85             std::size_t targetCharPos = 0;
86
87             for (std::size_t i = 0; i < loc.position; i++) {
88                 if (iss >> temp) {
89                     targetCharPos += temp.length() + 1;
90                 }
91             }
92
93             std::size_t start = targetCharPos > 20 ? targetCharPos - 20 : 0;
94             std::size_t length = std::min<static_cast<std::size_t>(40),
95                 context.length() - start>;
96             if (start < context.length()) {
97                 results.push_back({loc.line, context.substr(start, length)});
98             }
99         }
100     }
101
102     return results;
103 }
104
105 void bubbleSort(vector<pair<string, int>>& arr) {
106     for (std::size_t i = 0; i < arr.size(); i++) {
107         for (std::size_t j = 0; j < arr.size() - i - 1; j++) {
108             if (arr[j].second < arr[j + 1].second) {
109                 swap(arr[j], arr[j + 1]);
110             }
111         }
112     }
113 }
114
115 vector<pair<string, int>> getFrequentWords(int count) {

```

```

116     vector<pair<string, int>> wordFreq;
117
118     for (const auto& entry : index) {
119         if (entry.second.size() <= static_cast<std::size_t>(INT_MAX)) {
120             wordFreq.push_back({entry.first, static_cast<int>(entry.second.size())});
121         }
122     }
123
124     bubbleSort(wordFreq);
125
126     if (static_cast<int>(wordFreq.size()) > count) {
127         wordFreq.resize(static_cast<std::size_t>(count));
128     }
129
130     return wordFreq;
131 }
132
133 void addTrivialWord(const string& word) {
134     string cleanedWord = cleanWord(word);
135     if (!cleanedWord.empty()) {
136         trivialWords.insert(cleanedWord);
137         index.erase(cleanedWord);
138     }
139 }
140
141 void removeTrivialWord(const string& word) {
142     string cleanedWord = cleanWord(word);
143     trivialWords.erase(cleanedWord);
144     if (hasIndex()) {
145         buildIndex(currentFile);
146     }
147 }
148
149 bool hasIndex() const {
150     return !currentFile.empty();
151 }
152 };
153
154 int main() {
155     Concordance concordance;
156     string choice;
157
158     while (true) {
159         cout << "\nConcordance Menu:" << endl;
160         cout << "1. Build index from file" << endl;
161         cout << "2. Search word" << endl;
162         cout << "3. Show frequent words" << endl;
163         cout << "4. Add trivial word" << endl;
164         cout << "5. Remove trivial word" << endl;
165         cout << "6. Exit" << endl;
166         cout << "Enter choice: ";
167
168         getline(cin, choice);
169

```

```

170     if (choice == "1") {
171         string filename;
172         cout << "Enter filename: ";
173         getline(cin, filename);
174
175         ifstream test(filename);
176         if (test.good()) {
177             concordance.buildIndex(filename);
178             cout << "Index built successfully" << endl;
179         } else {
180             cout << "File not found" << endl;
181         }
182     }
183     else if (choice == "2") {
184         if (!concordance.hasIndex()) {
185             cout << "Please build index first" << endl;
186             continue;
187         }
188         string word;
189         cout << "Enter word to search: ";
190         getline(cin, word);
191
192         auto results = concordance.searchWord(word);
193         if (results.empty()) {
194             cout << "Word not found" << endl;
195         } else {
196             for (const auto& result : results) {
197                 cout << "Line " << result.first << ": ..." << result.second << endl;
198             }
199         }
200     }
201     else if (choice == "3") {
202         if (!concordance.hasIndex()) {
203             cout << "Please build index first" << endl;
204             continue;
205         }
206         cout << "How many top words to show? ";
207         int count;
208         cin >> count;
209         cin.ignore();
210
211         auto frequent = concordance.getFrequentWords(count);
212         for (const auto& word : frequent) {
213             cout << word.first << ": " << word.second << " occurrences" << endl;
214         }
215     }
216     else if (choice == "4") {
217         string word;
218         cout << "Enter trivial word to add: ";
219         getline(cin, word);
220         concordance.addTrivialWord(word);
221     }
222     else if (choice == "5") {
223         string word;

```

```

224         cout << "Enter trivial word to remove: ";
225         getline(cin, word);
226         concordance.removeTrivialWord(word);
227     }
228     else if (choice == "6") {
229         break;
230     }
231 }
232
233 return 0;
234 }

```

e_prevost@ares:~/Project_2\$ CPP concordance
concordance.cpp***

e_prevost@ares:~/Project_2\$./concordance.out

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 1
Enter filename: data.txt
Index built successfully

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 2
Enter word to search: declare
Line 1: ...es that they should declare the causes w...
Line 63: ...olemnly publish and declare, that these ...

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 3
How many top words to show? 5
of: 76 occurrences
and: 57 occurrences
for: 29 occurrences
our: 26 occurrences
has: 20 occurrences

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 4
Enter trivial word to add: of

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 3
How many top words to show? 4
and: 57 occurrences
for: 29 occurrences
our: 26 occurrences
has: 20 occurrences

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 4
Enter trivial word to add: and

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 5
Enter trivial word to remove: of

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 3
How many top words to show? 5

of: 76 occurrences
for: 29 occurrences
our: 26 occurrences
has: 20 occurrences
their: 20 occurrences

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 2
Enter word to search: independence
Word not found

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 2
Enter word to search: then
Word not found

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice:

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 2
Enter word to search: of
Line 1: ...When, in the course of human events, it ...
Line 1: ...e, among the powers of the earth, the se...
Line 1: ...n to which the laws of nature and of nat...
Line 1: ... laws of nature and of nature's God enti...
Line 1: ...ect to the opinions of mankind requires ...
Line 3: ...ty, and the pursuit of happiness.--That ...
Line 3: ...rs from the consent of the governed, tha...
Line 3: ...t whenever any form of government become...

Line 3: ...becomes destructive of these ends, it is...
Line 3: ...ds, it is the right of the people to alt...
Line 3: ...t when a long train of abuses and usurpa...
Line 3: ... patient sufferance of these Colonies; a...
Line 3: ...heir former systems of government. The h...
Line 3: ...rnment. The history of the present King ...
Line 3: ...of the present King of Great Britain is ...
Line 3: ...ritain is a history of repeated injuries...
Line 3: ...t the establishment of an absolute tyran...
Line 7: ...ernors to pass laws of immediate and pre...
Line 9: ...r the accommodation of large districts o...
Line 9: ... of large districts of people, unless th...
Line 9: ...elinquish the right of representation in...
Line 11: ...from the depository of their public reco...
Line 11: ...or the sole purpose of fatiguing them in...
Line 13: ...sions on the rights of the people.
...
Line 15: ...e powers, incapable of annihilation, hav...
Line 15: ... to all the dangers of invasion from wit...
Line 17: ...vent the population of these States; for...
Line 17: ... for naturalization of foreigners; refus...
Line 17: ...sing the conditions of new appropriation...
Line 17: ... new appropriations of lands.
...
Line 19: ... the administration of justice, by refus...
Line 21: ...one, for the tenure of their offices, an...
Line 21: ... amount and payment of their salaries.
...
Line 23: ...erected a multitude of new offices, and ...
Line 23: ... sent hither swarms of officers to haras...
Line 25: ... among us, in times of peace, standing a...
Line 25: ...without the consent of our legislatures...
Line 27: ...ilitary independent of and superior to t...
Line 29: ...ssent to their acts of pretended legisla...
Line 31: ...tering large bodies of armed troops amon...
Line 33: ... on the inhabitants of these States:
...
Line 35: ...rade with all parts of the world:
...
Line 39: ... us, in many cases, of the benefits of t...
Line 39: ...es, of the benefits of trial by jury:
...
Line 43: ...ing the free system of English laws in a...
Line 45: ...amentally the forms of our governments:
...
Line 49: ...by declaring us out of his protection, a...
Line 51: ...destroyed the lives of our people.
...
Line 53: ...orting large armies of foreign mercenari...
Line 53: ... complete the works of death, desolation...
Line 53: ... with circumstances of cruelty and perfi...
Line 53: ...y unworthy the head of a civilized natio...
Line 55: ...me the executioners of their friends and...
Line 57: ... on the inhabitants of our frontiers, th...

Line 57: ...s, whose known rule of warfare is an und...
Line 57: ...guished destruction of all ages, sexes, ...
Line 59: ...In every stage of these oppressions we h...
Line 59: ...fit to be the ruler of a free people.
...
Line 61: ... from time to time, of attempts by their...
Line 61: ... have reminded them of the circumstances...
Line 61: ...f the circumstances of our emigration an...
Line 61: ...ed them by the ties of our common kindre...
Line 61: ...n deaf to the voice of justice and of co...
Line 61: ...oice of justice and of consanguinity. We...
Line 61: ...as we hold the rest of mankind, enemies ...
Line 63: ...the Representatives of the United States...
Line 63: ...f the United States of America, in Gener...
Line 63: ...o the Supreme Judge of the world for the...
Line 63: ...d for the rectitude of our intentions, d...
Line 63: ...nd by the authority of the good people o...
Line 63: ... of the good people of these Colonies, s...
Line 63: ...d Colonies are, and of right ought to be...
Line 63: ... them and the State of Great Britain is,...
Line 63: ...ependent States may of right do. And for...
Line 63: ...And for the support of this declaration,...
Line 63: ...e on the protection of Divine Providence...

Concordance Menu:
1. Build index from file
2. Search word
3. Show frequent words
4. Add trivial word
5. Remove trivial word
6. Exit
Enter choice: 6
e_prevost@ares:~/Project_2\$ exit
exit

Script done on 2024-12-05 12:38:44-06:00 [COMMAND_EXIT_CODE="0"]