# **CSCI203 A1 Code Report**

Ethan Dakin | SN: 8209194

## **Code Description**

A brief analysis of my code is as follows:

My code first takes the input of the user, then reads the file. It reads character by character, comparing if it is a whitespace or not. If it is not a whitespace, then it checks if it is in the alphabet or not. If it is in the alphabet, then it adds the character to the string pool, and if it is not, then it checks if the next character is a space. If it is, then it increases the word counter. Following this, a for loop is initialized, which begins counting each word, with a complexity of O(n^2). Following this, I use quicksort to sort them from highest to lowest, and alphabetically if they are the same value. Then, my program prints these values to the console. Here is a simplified version of my program in pseudocode.

```
SIZE = 50000

WORD_SIZE = 20

text[SIZE * WORD_SIZE]

start[SIZE]

record word {
  word[WORD_SIZE],
  count
}

count[SIZE]

initialise
  input(filename)
```

```
character
index = 0
words = 0
character_index = 0
while (read character)
  if character is not whitespace then
   if character_index is 0 then
     start[words] = index
   fi
   if character is in the alphabet then
     text[index] = lowercase character
     index ++
   fi
   if next character is whitespace then
     words++
     character_index = 0
   else
     character_index++
   fi
  fi
elihw
```

```
count_index = 0
for i = 0 to words do
  begin = start[i]
 end = start[i + 1] - 1
  if i is equal to words - 1 then
    end = index - 1
  fi
 // null terminate
  temp[(end - begin) + 1] = endofstring
  found = -1
  for int j to words do
    if temp is equal to count[j].word then
     found = j
      break
    fi
  rof
  if found is greater than or equal to 0 then
    count[found].count++
  else
    copy_word(count[count_index].word, temp)
```

```
count[count_index].count = 1
   count_index++
 fi
rof
quicksort(count)
print "The first 10 words sorted alphabetically within frequency:"
for i is less than 10 do
  print "The word" + count[i].word + "occurs" + count[i].count + "times."
rof
print "The last 10 words sorted alphabetically within frequency:"
for i = count_index -1 and i is less than count_index do
  print "The word" + count[i].word + "occurs" + count[i].count + "times."
rof
print "The unique words sorted alphabetically:"
for i = 0 to count_index do
 if count[i].count == 1 then
   print "The word" + count[i].word + "occurs" + count[i].count + "times."
```

rof

esilaitini

#### **Complexity Analysis**

My code uses two main algorithms to sort itself.

Firstly, from line 109 to line 163, I get the count of each word. As I do not have a hash map, or use a data structure like a Heap, this complexity is unfortunately  $O(n^2)$ , as I first initialize a for loop from int i = 0 to n words; then initialize a for loop inside that code block, for the same condition, making it  $n^2$ .

Secondly, I use quicksort to sort numerically and alphabetically. As I simply change the if statement, this does not impact the sorting time.

Due to the dominant time complexity being  $O(n^2)$ , this is the time complexity of my program.

### **Data Structures**

In my code, I used a string pool for the character input, and an array of records for sorting.

As records are structs, I could essentially store a value for count and word in there, which, although not the access time is not as efficient as a dictionary / hash map, it helped reduce memory usage as I did not need two arrays of the same size.

For my string pool, I collected the input in by each character, and if it was whitespace or punctuation (and the next character was whitespace), it would add the current index to the start array, at the word index, and if it was an alphabetical character, it would add that to the text array.

#### Output

#### **Screenshot:**

```
PS D:\Projects\CSCI203> cd A1
  PS D:\Projects\CSCI203\A1> g++ ./main.cpp -o main;./main.exe Please enter the name of the input file:
    The first 10 words sorted alphabetically within frequency:
  The word: the occurs 131 times.
The word: of occurs 95 times.
The word: and occurs 94 times.
The word: to occurs 91 times.
    The word: you occurs 74 times.
The word: a occurs 64 times.
The word: i occurs 63 times.
    The word: he occurs 60 times.
    The word: was occurs 48 times. The word: mr occurs 47 times.
    The last 10 words sorted alphabetically within frequency:
   The word: window occurs 1 times.
The word: withdrew occurs 1 times.
The word: within occurs 1 times.
The word: without occurs 1 times.
    The word: wonderfully occurs 1 times. The word: wore occurs 1 times.
    The word: www.gutenbergorg occurs 1 times. The word: yet occurs 1 times.
    The word: yourself occurs 1 times.
   The unique words sorted alphabetically: The word: abuse occurs 1 times.
   The word: accomplished occurs 1 times.
The word: account occurs 1 times.
    The word: acknowledged occurs 1 times. The word: acquaintances occurs 1 times.
    The word: act occurs 1 times.
The word: added occurs 1 times.
  The word: added occurs 1 times. The word: addressed occurs 1 times. The word: adjusting occurs 1 times. The word: admiration occurs 1 times. The word admiration occurs 1 times. The word: admire occurs 1 times. The word: advice occurs 1 times. The word: advice occurs 1 times. The word: affect occurs 1 times. The word: affect occurs 1 times. The word: afterwards occurs 1 times. The word: afterwards occurs 1 times. The word: agree occurs 1 times.
   The word: agree occurs 1 times. The word: agreed occurs 1 times.
    The word: ah occurs 1 times.
The word: air occurs 1 times.
    The word: almost occurs 1 times.
The word: already occurs 1 times.
The word: altogether occurs 1 times.
   The word: altogethermr occurs 1 times. The word: amends occurs 1 times. The word: amiable occurs 1 times. The word: amiable occurs 1 times. The word: amongst occurs 1 times.
    The word: amusement occurs 1 times.
The word: angry occurs 1 times.
    The word: ankle occurs 1 times. The word: anyone occurs 1 times.
    The word: anywhere occurs 1 times. The word: arrival occurs 1 times.
The word: arrival occurs 1 times. The word: arrived occurs 1 times. The word: ascertaining occurs 1 times. The word: assemblies occurs 1 times. The word: assistance occurs 1 times. The word: assuring occurs 1 times. The word: astonishment occurs 1 times. The word: attacked occurs 1 times.
```

#### **Text Output:**

The first 10 words sorted alphabetically within frequency:

The word: the occurs 131 times.

The word: of occurs 95 times.

The word: and occurs 94 times.

The word: to occurs 91 times.

The word: you occurs 74 times.

The word: a occurs 64 times.

The word: i occurs 63 times.

The word: he occurs 60 times.

The word: was occurs 48 times.

The word: mr occurs 47 times.

The last 10 words sorted alphabetically within frequency:

The word: window occurs 1 times.

The word: withdrew occurs 1 times.

The word: within occurs 1 times.

The word: without occurs 1 times.

The word: wonderfully occurs 1 times.

The word: wore occurs 1 times.

The word: worth occurs 1 times.

The word: www.gutenbergorg occurs 1 times.

The word: yet occurs 1 times.

The word: yourself occurs 1 times.

The unique words sorted alphabetically:

The word: abuse occurs 1 times.

The word: accomplished occurs 1 times.

The word: account occurs 1 times.

The word: acknowledged occurs 1 times.

The word: acquaintances occurs 1 times.

The word: act occurs 1 times.

The word: added occurs 1 times.

The word: addressed occurs 1 times.

The word: adjusting occurs 1 times.

The word: admiration occurs 1 times.

The word: admire occurs 1 times.

The word: admitted occurs 1 times.

The word: advice occurs 1 times.

The word: affect occurs 1 times.

The word: afraid occurs 1 times.

The word: afterwards occurs 1 times.

The word: agree occurs 1 times.

The word: agreed occurs 1 times.

The word: ah occurs 1 times.

The word: air occurs 1 times.

The word: almost occurs 1 times.

The word: already occurs 1 times.

The word: altogether occurs 1 times.

The word: altogethermr occurs 1 times.

The word: amends occurs 1 times.

The word: amiable occurs 1 times.

The word: amongst occurs 1 times.

The word: amusement occurs 1 times.

The word: angry occurs 1 times.

The word: ankle occurs 1 times.

The word: anyone occurs 1 times.

The word: anywhere occurs 1 times.

The word: arrival occurs 1 times.

The word: arrived occurs 1 times.

The word: ascertaining occurs 1 times.

The word: assemblies occurs 1 times.

The word: assistance occurs 1 times.

The word: assuring occurs 1 times.

The word: astonishment occurs 1 times.

The word: attacked occurs 1 times.

The word: attention occurs 1 times.

The word: author occurs 1 times.

The word: away occurs 1 times.

The word: aye occurs 1 times.

The word: back occurs 1 times.

The word: barefaced occurs 1 times.

The word: behaviour occurs 1 times.

The word: beheld occurs 1 times.

The word: behind occurs 1 times.

The word: believe occurs 1 times.

The word: best occurs 1 times.

The word: bingleys occurs 1 times.

The word: bit occurs 1 times.

The word: bitterness occurs 1 times.

The word: black occurs 1 times.

The word: blue occurs 1 times.

The word: book occurs 1 times.

The word: books occurs 1 times.

The word: boulanger occurs 1 times.

The word: branch occurs 1 times.

The word: bring occurs 1 times.

The word: brotherinlaw occurs 1 times.

The word: brought occurs 1 times.

The word: called occurs 1 times.

The word: caprice occurs 1 times.

The word: care occurs 1 times.

The word: cases occurs 1 times.

The word: catching occurs 1 times.

The word: catherine occurs 1 times.

The word: certain occurs 1 times.

The word: chaise occurs 1 times.

The word: chance occurs 1 times.

The word: charming occurs 1 times.

The word: check occurs 1 times.

The word: children occurs 1 times.

The word: choose occurs 1 times.

The word: chooses occurs 1 times.

The word: circulation occurs 1 times.

The word: circumspection occurs 1 times.

The word: closed occurs 1 times.

The word: coat occurs 1 times.

The word: coldly occurs 1 times.

The word: comforted occurs 1 times.

The word: company occurs 1 times.

The word: compared occurs 1 times.

The word: conceited occurs 1 times.

The word: conjecturing occurs 1 times.

The word: consent occurs 1 times.

The word: consequence occurs 1 times.

The word: consequently occurs 1 times.

The word: consideration occurs 1 times.

The word: considered occurs 1 times.

The word: consisted occurs 1 times.

The word: contain occurs 1 times.

The word: continued occurs 1 times.

The word: contrast occurs 1 times.

The word: conversation occurs 1 times.

The word: copy occurs 1 times.

The word: cordial occurs 1 times.

The word: cost occurs 1 times.

The word: coughing occurs 1 times.

The word: coughs occurs 1 times.

The word: country occurs 1 times.

The word: courses occurs 1 times.

The word: cousin occurs 1 times.

The word: credit occurs 1 times.

The word: crown occurs 1 times.

The word: curiosity occurs 1 times.

The word: dances occurs 1 times.

The word: days occurs 1 times.

The word: deal occurs 1 times.

The word: declare occurs 1 times.

The word: declared occurs 1 times.

The word: decline occurs 1 times.

The word: declined occurs 1 times.

The word: deep occurs 1 times.

The word: deferred occurs 1 times.

The word: deigned occurs 1 times.

The word: delight occurs 1 times.

The word: derbyshire occurs 1 times.

The word: desire occurs 1 times.

The word: determined occurs 1 times.

The word: determining occurs 1 times.

The word: develop occurs 1 times.

The word: different occurs 1 times.

The word: difficult occurs 1 times.

The word: disappointed occurs 1 times.

The word: disclosed occurs 1 times.

The word: disconcerted occurs 1 times.

The word: discontented occurs 1 times.

The word: discovered occurs 1 times.

The word: discretion occurs 1 times.

The word: disgust occurs 1 times.

The word: dislike occurs 1 times.

The word: dispatched occurs 1 times.

The word: disposition occurs 1 times.

The word: distant occurs 1 times.

The word: distinguished occurs 1 times.

The word: dont occurs 1 times.

The word: door occurs 1 times.

The word: draw occurs 1 times.

The word: dresses occurs 1 times.

The word: drew occurs 1 times.

The word: during occurs 1 times.

The word: earliest occurs 1 times.

The word: early occurs 1 times.

The word: easy occurs 1 times.

The word: either occurs 1 times.

The word: elegant occurs 1 times.

The word: else occurs 1 times.

The word: eluded occurs 1 times.

The word: emphatic occurs 1 times.

The word: employed occurs 1 times.

The word: enduring occurs 1 times.

The word: engage occurs 1 times.

The word: engaged occurs 1 times.

The word: england occurs 1 times.

The word: enjoy occurs 1 times.

The word: enquired occurs 1 times.

The word: entering occurs 1 times.

The word: entrance occurs 1 times.

The word: equally occurs 1 times.

The word: escape occurs 1 times.

The word: establishment occurs 1 times.

The word: estate occurs 1 times.

The word: etc occurs 1 times.

The word: event occurs 1 times.

The word: exaggeration occurs 1 times.

The word: excessively occurs 1 times.

The word: exclamation occurs 1 times.

The word: expectations occurs 1 times.

The word: expected occurs 1 times.

The word: experience occurs 1 times.

The word: extracts occurs 1 times.

The word: extraordinary occurs 1 times.

The word: extremely occurs 1 times.

The word: eye occurs 1 times.

The word: fall occurs 1 times.

The word: falling occurs 1 times.

The word: families occurs 1 times.

The word: family occurs 1 times.

The word: fancied occurs 1 times.

The word: fancy occurs 1 times.

The word: fancying occurs 1 times.

The word: fashion occurs 1 times.

The word: fastidious occurs 1 times.

The word: fatigued occurs 1 times.

The word: favourable occurs 1 times.

The word: fear occurs 1 times.

The word: fears occurs 1 times.

The word: features occurs 1 times.

The word: felt occurs 1 times.

The word: fifth occurs 1 times.

The word: figure occurs 1 times.

The word: find occurs 1 times.

The word: finery occurs 1 times.

The word: fixed occurs 1 times.

The word: flatter occurs 1 times.

The word: flying occurs 1 times.

The word: fond occurs 1 times.

The word: forbidding occurs 1 times.

The word: forget occurs 1 times.

The word: forms occurs 1 times.

The word: fortnights occurs 1 times.

The word: fourth occurs 1 times.

The word: fretfully occurs 1 times.

The word: gave occurs 1 times.

The word: gentleman occurs 1 times.

The word: gentlemanlike occurs 1 times.

The word: given occurs 1 times.

The word: gods occurs 1 times.

The word: going occurs 1 times.

The word: goodhumoured occurs 1 times.

The word: goodlooking occurs 1 times.

The word: got occurs 1 times.

The word: gown occurs 1 times.

The word: gratified occurs 1 times.

The word: grieved occurs 1 times.

The word: grownup occurs 1 times.

The word: handsomer occurs 1 times.

The word: happily occurs 1 times.

The word: hat occurs 1 times.

The word: hate occurs 1 times.

The word: heart occurs 1 times.

The word: hearty occurs 1 times.

The word: heavens occurs 1 times.

The word: hertfordshire occurs 1 times.

The word: highly occurs 1 times.

The word: horrid occurs 1 times.

The word: horse occurs 1 times.

The word: house occurs 1 times.

The word: housekeeping occurs 1 times.

The word: hursts occurs 1 times.

The word: hypocritical occurs 1 times.

The word: idea occurs 1 times.

The word: ideas occurs 1 times.

The word: ignorant occurs 1 times.

The word: ill occurs 1 times.

The word: im occurs 1 times.

The word: imagine occurs 1 times.

The word: immediately occurs 1 times.

The word: included occurs 1 times.

The word: information occurs 1 times.

The word: ingenious occurs 1 times.

The word: inhabitants occurs 1 times.

The word: instead occurs 1 times.

The word: insufficient occurs 1 times.

The word: insupportable occurs 1 times.

The word: intelligence occurs 1 times.

The word: intended occurs 1 times.

The word: interrupted occurs 1 times.

The word: introduction occurs 1 times.

The word: its occurs 1 times.

The word: janes occurs 1 times.

The word: join occurs 1 times.

The word: joke occurs 1 times.

The word: joy occurs 1 times.

The word: keep occurs 1 times.

The word: king occurs 1 times.

The word: kingdom occurs 1 times.

The word: knowledge occurs 1 times.

The word: lace occurs 1 times.

The word: laid occurs 1 times.

The word: laws occurs 1 times.

The word: learnt occurs 1 times.

The word: least occurs 1 times.

The word: left occurs 1 times.

The word: less occurs 1 times.

The word: library occurs 1 times.

The word: license occurs 1 times.

The word: likely occurs 1 times.

The word: likes occurs 1 times.

The word: lines occurs 1 times.

The word: live occurs 1 times.

The word: lived occurs 1 times.

The word: london occurs 1 times.

The word: londonhis occurs 1 times.

The word: longbourn occurs 1 times.

The word: looking occurs 1 times.

The word: lose occurs 1 times.

The word: loved occurs 1 times.

The word: making occurs 1 times.

The word: mamma occurs 1 times.

The word: maria occurs 1 times.

The word: matter occurs 1 times.

The word: meaning occurs 1 times.

The word: meant occurs 1 times.

The word: meet occurs 1 times.

The word: mention occurs 1 times.

The word: mentioned occurs 1 times.

The word: met occurs 1 times.

The word: michaelmas occurs 1 times.

The word: mien occurs 1 times.

The word: might occurs 1 times.

The word: mind occurs 1 times.

The word: minds occurs 1 times.

The word: mistake occurs 1 times.

The word: mixture occurs 1 times.

The word: moment occurs 1 times.

The word: monday occurs 1 times.

The word: morris occurs 1 times.

The word: name occurs 1 times.

The word: near occurs 1 times.

The word: neglect occurs 1 times.

The word: neighbour occurs 1 times.

The word: nervous occurs 1 times.

The word: new occurs 1 times.

The word: newcomers occurs 1 times.

The word: news occurs 1 times.

The word: noble occurs 1 times.

The word: nobody occurs 1 times.

The word: none occurs 1 times.

The word: nor occurs 1 times.

The word: north occurs 1 times.

The word: number occurs 1 times.

The word: objection occurs 1 times.

The word: observing occurs 1 times.

The word: occasionally occurs 1 times.

The word: odd occurs 1 times.

The word: office occurs 1 times.

The word: often occurs 1 times.

The word: old occurs 1 times.

The word: online occurs 1 times.

The word: opinion occurs 1 times.

The word: out occurs 1 times.

The word: overscrupulous occurs 1 times.

The word: park occurs 1 times.

The word: part occurs 1 times.

The word: particular occurs 1 times.

The word: particularly occurs 1 times.

The word: passed occurs 1 times.

The word: people occurs 1 times.

The word: person occurs 1 times.

The word: persuade occurs 1 times.

The word: pieces occurs 1 times.

The word: planned occurs 1 times.

The word: playful occurs 1 times.

The word: pleasantly occurs 1 times.

The word: pleasing occurs 1 times.

The word: pleasure occurs 1 times.

The word: poor occurs 1 times.

The word: popularity occurs 1 times.

The word: preference occurs 1 times.

The word: press occurs 1 times.

The word: pretend occurs 1 times.

The word: promised occurs 1 times.

The word: pronounced occurs 1 times.

The word: property occurs 1 times.

The word: protested occurs 1 times.

The word: proud occurs 1 times.

The word: proudest occurs 1 times.

The word: punishment occurs 1 times.

The word: qualities occurs 1 times.

The word: questions occurs 1 times.

The word: quick occurs 1 times.

The word: quickness occurs 1 times.

The word: quieted occurs 1 times.

The word: quieter occurs 1 times.

The word: raised occurs 1 times.

The word: raptures occurs 1 times.

The word: rather occurs 1 times.

The word: read occurs 1 times.

The word: really occurs 1 times.

The word: recommend occurs 1 times.

The word: reflection occurs 1 times.

The word: regardless occurs 1 times.

The word: related occurs 1 times.

The word: remained occurs 1 times.

The word: reply occurs 1 times.

The word: resentfully occurs 1 times.

The word: resentment occurs 1 times.

The word: reserve occurs 1 times.

The word: respect occurs 1 times.

The word: restrictions occurs 1 times.

The word: reuse occurs 1 times.

The word: ridiculous occurs 1 times.

The word: rightful occurs 1 times.

The word: rode occurs 1 times.

The word: round occurs 1 times.

The word: rudeness occurs 1 times.

The word: sakes occurs 1 times.

The word: sarcastic occurs 1 times.

The word: sat occurs 1 times.

The word: satisfactory occurs 1 times.

The word: save occurs 1 times.

The word: says occurs 1 times.

The word: scarcity occurs 1 times.

The word: scolding occurs 1 times.

The word: secondhand occurs 1 times.

The word: seek occurs 1 times.

The word: seemed occurs 1 times.

The word: seen occurs 1 times.

The word: selfish occurs 1 times.

The word: sensible occurs 1 times.

The word: servants occurs 1 times.

The word: serving occurs 1 times.

The word: setdowns occurs 1 times.

The word: settling occurs 1 times.

The word: seven occurs 1 times.

The word: several occurs 1 times.

The word: share occurs 1 times.

The word: sharpened occurs 1 times.

The word: shocking occurs 1 times.

The word: shut occurs 1 times.

The word: sick occurs 1 times.

The word: sight occurs 1 times.

The word: silly occurs 1 times.

The word: sit occurs 1 times.

The word: sitting occurs 1 times.

The word: six occurs 1 times.

The word: sixth occurs 1 times.

The word: skill occurs 1 times.

The word: smiles occurs 1 times.

The word: solace occurs 1 times.

The word: somebody occurs 1 times.

The word: somewhat occurs 1 times.

The word: sorry occurs 1 times.

The word: speak occurs 1 times.

The word: speaking occurs 1 times.

The word: spirits occurs 1 times.

The word: splendid occurs 1 times.

The word: spoke occurs 1 times.

The word: sprained occurs 1 times.

The word: stared occurs 1 times.

The word: starting occurs 1 times.

The word: step occurs 1 times.

The word: stoutly occurs 1 times.

The word: stranger occurs 1 times.

The word: stress occurs 1 times.

The word: struck occurs 1 times.

The word: stupid occurs 1 times.

The word: suddenly occurs 1 times.

The word: suffer occurs 1 times.

The word: sufficient occurs 1 times.

The word: suiting occurs 1 times.

The word: suppositions occurs 1 times.

The word: surely occurs 1 times.

The word: surmises occurs 1 times.

The word: surpassing occurs 1 times.

The word: surrounding occurs 1 times.

The word: talk occurs 1 times.

The word: talked occurs 1 times.

The word: tall occurs 1 times.

The word: tallest occurs 1 times.

The word: tear occurs 1 times.

The word: teasing occurs 1 times.

The word: temper occurs 1 times.

The word: tempt occurs 1 times.

The word: terms occurs 1 times.

The word: these occurs 1 times.

The word: third occurs 1 times.

The word: those occurs 1 times.

The word: thought occurs 1 times.

The word: threeandtwenty occurs 1 times.

The word: throw occurs 1 times.

The word: tide occurs 1 times.

The word: times occurs 1 times.

The word: tiresome occurs 1 times.

The word: title occurs 1 times.

The word: tolerable occurs 1 times.

The word: tomorrow occurs 1 times.

The word: toward occurs 1 times.

The word: towards occurs 1 times.

The word: trimming occurs 1 times.

The word: tumult occurs 1 times.

The word: turned occurs 1 times.

The word: turning occurs 1 times.

The word: unaffected occurs 1 times.

The word: uncertain occurs 1 times.

The word: uncommonly occurs 1 times.

The word: under occurs 1 times.

The word: understand occurs 1 times.

The word: understanding occurs 1 times.

The word: universally occurs 1 times.

The word: unless occurs 1 times.

The word: unlucky occurs 1 times.

The word: unreserved occurs 1 times.

The word: unworthy occurs 1 times.

The word: upper occurs 1 times.

The word: using occurs 1 times.

The word: various occurs 1 times.

The word: venture occurs 1 times.

The word: vexed occurs 1 times.

The word: vexing occurs 1 times.

The word: village occurs 1 times.

The word: violent occurs 1 times.

The word: visiting occurs 1 times.

The word: waited occurs 1 times.

The word: walking occurs 1 times.

The word: wasting occurs 1 times.

The word: wayswith occurs 1 times.

The word: week occurs 1 times.

The word: whatsoever occurs 1 times.

The word: whichever occurs 1 times.

The word: whom occurs 1 times.

The word: wifes occurs 1 times.

The word: window occurs 1 times.

The word: withdrew occurs 1 times.

The word: within occurs 1 times.

The word: without occurs 1 times.

The word: wonderfully occurs 1 times.

The word: wore occurs 1 times.

The word: worth occurs 1 times.

The word: www.gutenbergorg occurs 1 times.

The word: yet occurs 1 times.

The word: yourself occurs 1 times.