

COMP20010



Data Structures and Algorithms I

11 - Tutorial: Hashtables

Dr. Aonghus Lawlor
aonghus.lawlor@ucd.ie



Hashtables

- Write a Java program which uses `java.util.HashMap` to count the words in a file.
- Use the words from this file: https://csmoodle.ucd.ie/moodle/pluginfile.php/106305/mod_resource/content/1/sample_text.txt
- Report the top 10 most frequently used words.

Hashtable

- Implement the original Java String hashCode function:
- What is the hash code for the string: "java"
- use the ascii table
- <http://www.asciitable.com/>

```
public int hashCode() {  
    int hash = 0;  
    int skip = Math.max(1, length() / 8);  
    for (int i = 0; i < length(); i += skip)  
        hash = (hash * 37) + charAt(i);  
    return hash;  
}
```

Hashtable

- Using the list of words from
- https://csmoodle.ucd.ie/moodle/pluginfile.php/106306/mod_resource/content/1/words.txt
- Find the number of collisions using polynomial accumulation with $a=41$ (Ans: 4)
- Find the number of collisions using polynomial accumulation with $a=17$ (Ans: 387)
- Find the number of collisions using a cyclic shift with a shift value of 7 (Ans: 290)
- Find the number of collisions which occur using the old Java hash code function (Ans: 147)
- Comment on the number of collisions for different hash functions